University of Huddersfield Repository

Varley, Matthew

A Case Study on Understanding the Influence of Diegetic Audio on Immersion in a Third-Person Role-Playing Game

Original Citation

Varley, Matthew (2018) A Case Study on Understanding the Influence of Diegetic Audio on Immersion in a Third-Person Role-Playing Game. Masters thesis, University of Huddersfield.

This version is available at http://eprints.hud.ac.uk/id/eprint/34674/

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

http://eprints.hud.ac.uk/

A Case Study on Understanding the Influence of Diegetic Audio on Immersion in a Third-Person Role-Playing Game

Matthew Varley

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of MA by Research

The University of Huddersfield

January 2018

Abstract

This thesis investigates the influence of diegetic audio in relation to creating immersive gameplay experiences in third-person perspective games. Previous research suggests that immersion is key to an excellent gaming experience (Brown & Cairns, 2004). In particular, certain works show the influence of audio in relation to immersive experiences (Stockburger, 2003; Jorgensen, 2008, 2011; Grimshaw, 2007, 2012; Huiberts, 2010; Usher, 2012). Drawing on this evidence, I created a video-game wherein diegetic audio could be manipulated adaptively in an experimental environment as the game is played.

A convergent parallel mixed-method case study (Creswell, 2014) was developed to determine if removing, modifying, or de-synchronising elements of the diegetic soundscape in a third-person role-playing game (RPG) at runtime would negatively affect immersion in the game. First, I examined quantitative aspects of gamer preference, genre expectations, and previous experiences of immersion, to determine the likelihood of becoming immersed in the test game. Next, ten participants played the game whilst, unknown to them, auditory aspects of the game were modified. Finally, players took part in a semi-structured interview to reveal how they felt the audio changes influenced their immersion in the game, or if they noticed the changes.

Findings showed that participants were engaged sufficiently that none noticed the adaptation of key avatar sounds, such as footsteps, and none noticed a discontinuity in ambient background sounds. Players revealed that much of their attention was taken up with completing game objectives and learning the game controls. This suggests that when a player's attention is diverted by an objective, less attention is free to give to sounds that are not crucial for that task: challenge-based immersion is increased, whilst immersion in the sensory aspects of the game is decreased. Developing an experimental game environment offers further insight into the role of diegetic audio in third-person perspective games, and how it influences immersive gameplay experiences.

Copyright Statement

The following notes on copyright and the ownership of intellectual property rights must be included as written below:

i. The author of this thesis (including any appendices and/ or schedules to this thesis) owns any copyright in it (the "Copyright") and s/he has given The University of Huddersfield the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

ii. Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the University Library. Details of these regulations may be obtained from the Librarian. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

iii. The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

Acknowledgements

This thesis exists only because of the continued support, love, and unrelenting patience of my friends, family, and better half, Kristine. Your support and motivation throughout the last year and a half has been my strength and without it I would not be where I am now.

I would also like to thank my supervisor, Dr. Elizabeth Dobson, who's faith, encouragement, and guidance pushed me on through the difficult decisions and nail-biting moments that arose during the length of this thesis.

This work is dedicated to my bunny, Cookie, who had to be put to sleep in November 2017 at just six-months old, and who I will always feel a connection to through this work.

Table of Contents

List of Abbreviations

EQ = Equalisation.

GEQ = Game Engagement Questionnaire (Brockmyer, et al., 2009);

IEZA = Interface, effect, zone, affect; model of game-audio, Huiberts (2010).

ITQ = Immersive Tendencies Questionnaire; Witmer & Singer (1998).

NPC = Non-player character.

PUGS = “Pretty Ugly Game Sound” questionnaire; Huiberts (2010).

SCI = Sensory, challenge-based, imaginative; model of immersion, Ermi & Mara (2005).

SFX = Sound effects.

VR = Virtual reality.

UI = User-interface.

UX = User experience.

List of Figures

List of Tables

1. Introduction

The following research explores diegetic audio in third-person video-games, and how it can influence immersive gameplay experiences. I will consider the preconditions for immersion within video-games, and explore the role that diegetic audio has to play in this process as both a usability system and a means of encouraging presence in a virtual world. This will guide my decisions when developing an experimental game, one that features controls which enable the researcher to remove, modify, or de-synchronise elements of diegetic audio in the game-world at runtime. The aim is to see if modifying diegetic audio will cause the player to feel less immersed.

Immersion is a topic that spans many fields, and is not limited to video-game theory. Previous research focusing on the immersive experiences of video-games has used the first-person shooter genre as an example (Nacke & Lindley, 2008; Grimshaw, 2007, 2008; Stevens & Raybould, 2015), due to its ease of relatability and lack of avatar-related incoherence. For RPG games, too, the first-person point of view (POV) has been shown to be more immersive, as it allows players to project their thoughts and actions into the game-world (Denisova & Cairns, 2015). Some disagree, and instead believe that the extra perceptual affordability of the third-person POV builds a richer visual presentation of the game world, and brings more value for games requiring exploration and interaction (Taylor, 2002). Indeed, many of the most successful video-game franchises have been presented in the third-person, such as *Grand Theft Auto V* (Rockstar North, 2013), *Tomb Raider* (Core Design, 1996)*, Middle-earth: Shadow of*

Mordor (Monolith Productions, 2014)*, World of Warcraft* (Blizzard Entertainment, 2005), and the majority of massively multiplayer online role-playing games (Black, 2017). Regardless of perspective, research has highlighted the importance of creating immersive gameplay experiences for players, as well as the importance of creating testable definitions and measures for observing immersion for researchers (Nacke & Lindley, 2008).

Though the focus of this thesis is from an auditory perspective, it is crucial in the understanding of immersion that sound cannot be examined in isolation (O'Keeffe, 2011); other modalities of gameplay, such as visuals and game mechanics, must also be taken into consideration, as well as the varying classifications and definitions of immersion as it manifests itself in gameplay. In this thesis, I will draw on previous studies of audio-based immersion in video-games (Jorgensen, 2008; Huiberts, 2010; Usher, 2012), and general theories of immersion (Brown & Cairns, 2004; Ermi & Mayra, 2005; Cheng & Cairns, 2005), and apply them in the development of an experimental video-game and methodology that can be used to examine immersion in a third-person RPG from an auditory perspective.

2. Literature Review

The term immersion is used in relation to many subjects, and is not limited to the subject area of gaming, as Arsenhault (2005) summarises beautifully:

> Immersion occurs when one gazes at a painting, listens to music, is lost in a book or absorbed in a game of chess, so much that he ceases to perceive the museum or the sounds of the street, forgets the events happening in the real world, and suspends his knowledge of its laws.

Mediated environments have been considered 'immersive' long before the technological boom of the 20th century (Moser & Macleod, 1996). In early examples of theatre, both actors and audiences were engrossed in the narrative world; captivated, and swept into new dimensions (Frasca, 2001). Books and other narrative-based literature have also been found to exhibit in their readership examples of immersive tendencies (Biocca & Levy, 1995; Laramée, 2002). Television and film undoubtedly share the properties explored in this thesis, with the Cinerama of the early 1950's basing its entire marketing campaign on "immersing the viewer in different worlds" (Lombard & Ditton, 1997; Ijsselsteijn et al., 2000); and now, in the 21st century, video-games are taking centre stage as the newest forms of immersive media (Laramée, 2002).

From a psychological view, the playing of video-games is a complex blend of exploration, narrative, audio-visual stimuli, and user-input, that can elicit a dynamic range of cognitive activity. Video-games often require the player to make selections, interpret stimuli, and make choices to influence outcomes within the environment of the game-world; all of which encourages the player to invest more time and effort into the game, ultimately resulting in a relationship between the player and the symbol they control (Kline, Witherford, & de Peuter, 2003). As such, video-games present an ideal platform for observing immersion; the increased mental activity and continued need for focus restricts any cognitive processes that may otherwise be used to doubt the reality of the displayed environment (McMahan, 2003; Wissmath, Weibel, & Groner, 2009). However, issues have arisen in the past from the definitional issues that surround immersion, due to overlap between concepts such as presence and flow. It is therefore necessary to work through these previous classifications and definitions, to understand exactly what is being referred to when a game is described as being immersive.

2.1 Classifying, Supporting, and Disrupting Immersion

A conclusive definition of immersion is often disputed (Brown & Cairns, 2004; Collins, 2008), though it is a term often used by gamers, designers, reviewers, and researchers when describing gameplay experiences (Jennet et al., 2008; Ortqvist & Liljedahl, 2010). When immersed, individuals exhibit trance-like behaviours (Radford, 2000); they often lose track of time (Czsentmihalyi, 1991; Jennet, et al., 2008), perform actions unconsciously (Brown & Cairns, 2004), feel transported into the game-world (Ijsselsteijn, Freeman, & de Riddler, 2001), and identify or empathise with the situation the character is currently in (Huiberts, 2010). Though there is no "one-size-fits-all" definition of immersion in the gaming community (Kramer, 2016), immersion has previously been defined as a psychological state of heightened focus on a task or activity that provides a continual stream of cognitive stimulus (Witmer & Singer, 1998), and "the degree to which a virtual environment submerges the perceptual system of the user… the senses are immersed in the virtual world (Biocca & Delaney, 1995). As arguments continue around defining immersion, there arises another pertinent issue as to how the varying stages of immersion can be classified.

2.1.1 Immersive Classifications

There have been numerous attempts to classify and understand immersion. McMahan (2003) defined two ways of categorising immersion: perceptual immersion, and psychological immersion. Perceptual immersion is achieved by restricting the senses access to the outside world, thereby focusing all attention on the artificial world, and psychological immersion is the user's mental absorption within the game-world. To achieve a sense of psychological immersion in a 3-D computer game, the expectations of the user in regard to the game environment must match the conventions of the game environment; the actions of the user must have a noticeable, though not trivial, impact on the environment; and, lastly, the conventions of the game-world must remain consistent. Consistency is especially important, as, from a temporal perspective, video-games are a type of non-linear media.

Brown and Cairns (2004) examined the temporal nature of immersion as it unfolds over time whilst mediated by internal and external factors relating to gameplay. They reduced the immersive process to three stages: engagement, categorised by the player investing time and effort into playing; engrossment, which relates to features of game construction such as visuals, audio, and storyline; and total immersion, categorised by feelings of presence in the game-world and high levels of emotional involvement with the narrative, character, and game-world. As sound is also temporal, it is important to bare this work in mind when examining the immersive properties of sound (Jorgenson, 2011). Total immersion refers to feelings of presence, a separate theoretical construct to immersion but one that is frequently discussed in relation to virtual environments and video-games.

Presence is defined by Patrick et al. (2000, p.479) as: “the extent to which a person’s cognitive and perceptual systems are tricked into believing they are somewhere other than their physical location.” In the context of video-games, Madigan (2010) discusses the elements of a game construction that facilitate the player feeling spatially located in the game-world, and devises two groups of characteristics: “those that create a rich mental model of the game environment and those that create consistency between the things in that environment”. Environmental richness can be achieved through the presence of multiple channels of sensory information, and how complete and detailed the sensory information is. Environments that are cognitively demanding and present room for exploration and interpretation, and a well-devised and interesting narrative, also contribute to environmental richness. Another immersive theory that refers to these elements of game construction is the SCI model developed by Ermi and Mayra (2005).

The SCI (sensory, challenge-based, imaginative) model places aspects of immersion into three dimensions: sensory, imaginative, and challenge-based immersion. Sensory immersion arises from the relationship between visual and auditory aspects of the game, resulting in a consistent and believable game-world that engages the players senses as though they were present in the game-world (Ermi & Mayra, 2005). This bares similarities to Calleja's (2011) concept of incorporation, in that both reference the player being part of the game's acoustic world. Though sensory immersion is defined as being crucial, it is suggested that its role is a supportive one: supporting imaginative and challenge-based immersion. It is also believed that classifications of immersion can conflict during play, and that being immersed from a challenge-based or imaginative perspective can cause a decrease in sensory immersion (Westerberg & Schoenau-Fog, 2015).

Challenge-based immersion relates to in-game challenges and objectives, interaction, and fluid 'sensormotor abilities' such as adept controller use, as well as social interaction, cooperation, and competition, with other players (Ermi & Mayra, 2005). This shares similarities to Flow theory, created by Michael Csikszentmihalyi (1991), in that both describe a balance between skill and challenge that results in feelings of enjoyment when an individual is performing a rewarding activity that is of similar difficulty to their current level of competency, see Fig 1. Csikszentmihalyi (1991) describes eight key variables that define the flow state:

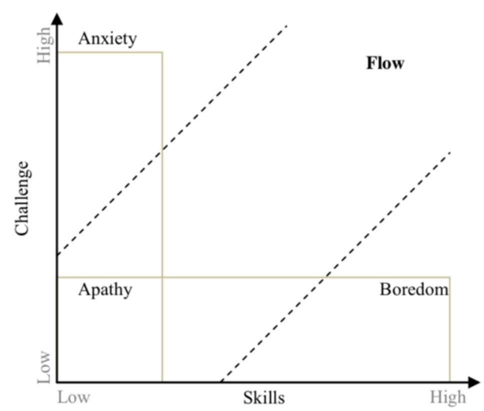

Figure 1: Conceptualising the flow state (Nacke & Lindley, 2008)

- Loss of self-consciousness
- Clear and achievable goals
- Distorted sense of time
- Feelings of enjoyment
- A balance between challenge skill
- Direct, immediate feedback
- Continual attention
- High degree of control

Examples of flow states are often found in relation to video-games, and gamers often describe the pleasure in playing a game they are good at, but also describe increases in difficulty level as pleasurable, as it provides a more challenging, and arguably rewarding, experience (Sherry, 2004; Hoffman & Novak, 2009; Jin, 2012). Finally, imaginative immersion relates to the emotional level of engagement and investment with the games' characters, narrative, and environment. Though Arsenault (2005) argues that this term is too vague and should be replaced with "Fictional immersion", as it reduces the broadness of the term "imagination" yet retains enough to include narration and representation that is commonly found within video-games. Classifications of immersion are necessary in order to fully understand the varying effects it has on players of video-games. Following this, it is now important to understand how immersion is supported and encouraged, from the perspective of both the games' design and the person playing it.

2.1.2 Supporting Immersion

Immersion can be supported by both internal and external factors of gameplay (Huiberts, 2010). Internal elements refer to a games' design, mechanics, controls, visuals, and audio; whilst external factors include: the physical environment, distractions, level of comfort, mood, personality, and the technology being used to play the game. However, total immersion is not achievable through the presence of just one factor, it is the combination and merging of numerous game and play-related factors, or modalities, that create the desired result (Cheng & Cairns, 2005). For a game designer, creating a sense of immersion depends on the type of game being designed and the artistic intention and features of the game (Collins, 2008; Madigan, 2010). Specific examples of such modalities include: graphical realism and suitability, behavioural realism, and high levels of user-environment interactivity.

Considering that the player's interaction with the game-world takes place through the controller, the naturalness and responsiveness of the game controls plays a key role in the immersive process, enabling the controller to become an extension of the players own senses within the virtual space of the game-world (McGloin, Farrar, & Krcmar, 2013; Kramer, 2016). Tasks or objectives to complete provide cognitive stimulation and facilitate challenge-based immersion and flow states (Csikszentmihalyi, 1991; Patrick, et al., 2000; Wissmath, Weibel, & Groner, 2009). Rich and detailed environments provide sensory immersion and encourage spatial presence (Marks, 2009). Narrative forms are well-known for their immersive properties and have been explored extensively in the past (Douglas & Hargadon, 2000; Ryan, 2001; Martinez, 2014). If a video-game narrative is to be considered immersive, it must flow, provide clear progression, have a defined pace, and provide opportunities to connect with characters; Modern games provide all this in novel-length tales and include the chance to directly interact with, and affect the outcome of, the story, making them highly engaging (Qin, Patrick Rau, & Salvendy, 2009; Lebowitz & Klug, 2011). Character identification is also defined as crucial for immersion, as the player must react to events in the game-world as they would in real life (Grimshaw, 2012). A final key point is consistency; game-worlds do not need to be audio-visually realistic (McMahan, 2003), in fact some games may benefit from the opposite (Madigan, 2010), however, a reduced, perceptual realism may be all that is required to form the basis from which immersion can work (Grimshaw, 2007). The game world plays a key role in creating an immersive experience, however, the environment of the user can also contribute.

From the perspective of a gamer, making their experience more perceptually immersive may involve using headphones to isolate all background noise, or speakers to feel the rumble of explosions (Huiberts, 2010; DiGiose, 2016); Multiple screens can also be used, or large projections screens, so that details within the game are easily visible (Laird, 2015; Porter, 2017), this has been shown to create a more immersive environment (Hou, et al., 2012); finally, customised game controllers allow users to define their own controls, for a natural and immersive experience (Brown, et al., 2011; Andronico, 2016). However, it must be noted that immersion is not totally dependent on the technology used to play (McMahan, 2003), the player must be willing to play and willing to let themselves fall into the immersive trance (Radford, 2000; Brown & Cairns, 2004). Personality traits of the individual can also affect their susceptibility to immersion and spatial transportation into mediated environments (Wissmath, Weibel, & Groner, 2009). The Immersive Tendencies Questionnaire (ITQ), developed by Witmer &

Singer (1998), enables researchers to reveal these personality traits and help determine an individual's likelihood of becoming immersed in an activity. Once the elements that encourage immersion have been examined, it is possible to begin investigating what happens when these elements are not present or are removed during play.

2.1.3 Disrupting Immersion

The research of Brown and Cairns (2004) defined three stages to immersion, which were explored in the previous chapter, though access to each of these levels is mediated by barriers, defined as: access, investment, game construction, empathy, and atmosphere, see Fig 2. Access refers to the genre-preference of the gamer and how comfortable they feel with the controls of the game; investment refers to the time and effort that the gamer invests in playing; game construction refers to how well the game is made, and includes visual quality, interesting tasks, and plot; empathy is defined as the growth of attachment to the game-world and character; and, finally, atmosphere, which is defined as the development of game construction relevant to the actions and locations of the characters. If the requirements of the barriers are not fulfilled, an individual cannot progress through to the final stage of total immersion. These barriers proved especially useful during the design phase of the game and were used as a

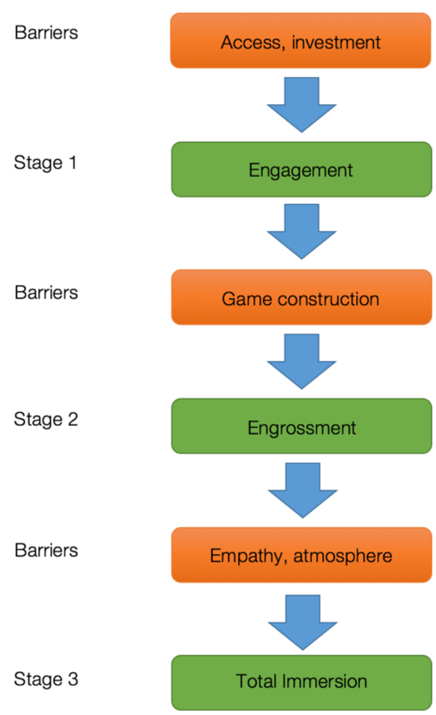

Figure 2: A diagram conceptualising the temporality of immersion

checklist to ensure that the game would present as many opportunities as possible for the player to pass through the barriers towards total immersion. Other research used deleterious usability to monitor how making changes to physical and control-based elements within a game during game-play can affect immersion.

Preventing efficient user interaction in a video-game can result in a less immersive experience (Roden, Parberry, & Ducret, 2005). However, the work of Cheng and Cairns (2005) showed that even in a simple, task-based game environment, players can be immersed to such a degree that they become unaware of any changes made within the game. 14 participants with backgrounds in computer science, and who regularly reported playing video-games, were asked to navigate a simple game-world to complete a task, see Fig. 3 for a map of the game-world. As they crossed the halfway line, shown by the red dotted line in Fig. 3, physical and visual qualities of the game were changed. Physical changes included changing the height the character is able to jump, and visual changes involved changing textures painted within the scene. These changes needed to be subtle, so that the experience of gameplay remained consistent and any breaks in immersion "could be safely attributed to the change in coherence rather than to any inherent bizarreness" (2005, p.1272). To the surprise of the authors, the results were conclusive that not one of the participants noticed any change in gameplay, such was the extent of their immersive state.

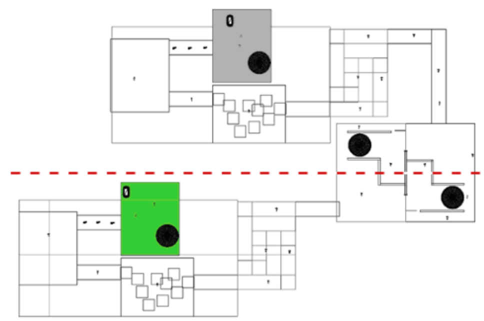

Figure 3: A diagram of the map used by Cheng and Cairns.

Huiberts (2010) conducted the PUGS ("Pretty Ugly Game Sound") questionnaire, in which participants were asked to recall examples of audio in games that disrupted their immersion and overall experience of the game. One question in particular, "Have you ever experienced that your feeling of immersion disappeared while you were playing a game? What was the reason?", highlighted two types of barriers to immersion: internal and external. Responses and categorisations can be seen below in Table 1.

Internal - barriers present in the game	**Causes**
Interruption of game play The pace of flow of the game is altered	- Non-interactive movements - Level completed - Changed mode of gameplay - Loading screens - Cut scenes - Game crash, bugs
Game construction Mistakes or unlikely instances in the game	- Behaviour of artificial intelligence - Weird sounds - Non-responsive controls - Unrealistic or unnatural things
Frustration of the player The player is annoyed by something in the game. This often is a result of the two other categories (interruption or construction) but not at all times.	- Frustration, irritation, or boredom - In-game advertisement - Getting stuck, unable to progress/succeed
External – barriers that disturb attention	**Causes**
Distraction (attention of the player)	- External sounds in the user environment - Other people (parents, partners, phones)
State (Mood or physical state of the player)	- Not in the mood / personal stress - Becoming exhausted - Pain, RSI, hunger

Table 1: The internal and external barriers to audio immersion (Huiberts, 2010)

2.1.4 Summary

When immersed, people exhibit trance-like behaviours, perform actions unconsciously, feel transported into the game-world, and identify with the games' character. Consequences of immersion are also viewable in the concepts of flow and presence, and as such this has led to difficulty in defining immersion in the context of video-games (McMahan, 2003; Collins, 2008). Immersion has been classified extensively, and to varying degrees of detail. Ermi and Mayra (2005) provide a three-dimensional summary of immersion: sensory, challenge-based, and imaginative immersion, the definitions of which will be used throughout the remainder of this work. Both the games' design and the player can heighten immersion; presenting the player with an in-game task (Patrick, et al., 2000), detailed and consistent game-worlds (Calleja, 2011), engaging characters and narrative (Douglas & Hargadon, 2000), and responsive controls (McGloin, Farrar, & Krcmar, 2013), are a few examples of game features that can increase feelings of immersion. Players can increase their own likelihood of immersion using large screens or projections, headphones, and customised input controllers (Porter, 2017; DiGiose, 2016; Andronico, 2016); though some individuals are more prone to immersion than others (Wissmath, Weibel, & Groner, 2009). From a temporal perspective, the path to immersion is mediated by three stages: engagement, engrossment, and total immersion; with access to these stages controlled by barriers relating to both internal and external game factors (Brown & Cairns, 2004; Huiberts, 2010). The information discussed above will prove crucial in the development of the game-world, the physical test environment, and the equipment used during the experiment, ensuring that it is all as immersive as possible. To understand more about how sound in particular can influence immersion, perceptual systems of hearing must be explored, as well as the differing classifications and functions of video-game audio, and how sound designers can create and implement sounds to achieve maximum immersion. Both the studies of Brown & Cairns (2004) and Cheng & Cairns (2005) share the view that once a certain level of immersion has already been achieved, incoherence in modalities becomes less of a distracting factor.

2.2 Perceptual Systems and Video-Game Audio

> Listen to the sounds involved… the sheer amount of sound effects involved is incredible…listen to the fabric of the clothes rubbing, the exhale of breath as he climbs objects, his interaction with the world… the size of his feet dictates the sound of his shoes, the textures of the ground he's climbing upon, the wind, the water lapping, the bus creaking, the birds tweeting, branches swaying, insects, flags fluttering… Despite not prominently hearing every individual sound that has been recorded in the mix, without them, it would just severely affect the immersion. Perryman (2016).

In the past, video-game audio has been underrated, and has played a lesser role in the gameplay experience, due, in part, to limitations of technology available at the time (Ekman, 2005; O'Keeffe, 2011). However, with advancements in both game design software and the platforms on which games are played, the role of audio has increased dramatically; though difficulties can sometimes arise during the integration of game audio due to the non-linearity of gameplay and possibility of multiple outcomes (Geelen, 2008; Westerberg & Schoenau-Fog, 2015). Difficulties aside, the role of audio in relation to creating immersive game experiences is well known; For a player to be successful in a video-game environment, they must perceive the game-world in 360 degrees, with the majority of information about the off-screen environment being transferred aurally (Morris, 2002; Whalen, 2004). But what happens if sounds are unsuitable, improperly synchronised, or completely missing from the game? If the function of the footstep sound is to locate the player in the game-world, what happens when these sounds are changed? Some insights are offered when considering the categorisation and functions of video-game audio, what is known about how sound encourages immersion in gameplay, and how the perceptual systems of listening decode sound signals into meaningful messages.

2.2.1 Hearing VS Listening

Hearing is the fastest of the five senses, though it is mainly a passive event (Horowitz, 2012; Roux-Girard, 2014). We cannot close our ears as we do our eyes, so our brain must filter sound appropriately. This allows sound designers and composers direct access to the subconscious (Sonnenschein, 2001). Listening, however, is an activity that has multiple modes and styles (Chion, 1994; Margulis, 2007), and unlike the visual aspects of a video-game, “listening is not a simulated experience” (O’Keeffe, 2011, p.51). Chion (1994) proposed a flexible framework for sound design that was rooted in film-sound theory, drawing heavily on Pierre Schaeffer’s analyses of sound objects. Chion divides listening into three modes: ‘causal listening’, ‘semantic listening’, and ‘reduced listening’. Causal listening is when the listener tries to discern the source of a sound and attempts to understand its cause. For example, sound can be used to predict upcoming events (Meyer, 1959; Huron, 2014), and to understand the environment around the listener from a pragmatic view (Liljedahl, 2011), such as for navigation and orientation (Grimshaw & Schott, 2008). Semantic listening is used by the listener to interpret the meaning of a sound, often using a culturally-defined semiotic code, and can help the listener derive meaning and context for a sound (Grimshaw, 2007). The final mode, reduced listening, is used less frequently by listeners, especially those playing video-games (Stockburger, 2003); it is used when listening to the sonic qualities of a sound without thinking about its origin or purpose. These perceptual systems have evolved to gather information from natural ecologies in the world around us, as “sound is an integral part of our everyday lives” (Liljedahl, 2011). In video-games, and cinema, the senses of the listener will interpret any auditory data as if it was being perceived in a natural ecology (Anderson, 1996; Grimshaw, 2007, 2012). Within the soundscape of the game-world, it is important to understand the different sound types, their functions, and how they can be categorised.

2.2.2 Sound Types and Categories

There is a vagueness surrounding the term video-game audio and what exactly it refers to in academic research. Huiberts (2010, p.21) clarifies this using the terms interactive and non-interactive audio, for audio heard during interactive gameplay, and audio during non-interactive moments, and anything outside of the game context, such as adverts. From this point, the term ‘interactive’ video-game audio will be used to clarify this distinction.

The way in which game-audio is categorised can depend largely on the game itself; its style and aesthetic functions. Kamp (2012, p.240) summarises that the soundscape of an abstract game, like *Tetris*, a game that has no diegesis at all, is inherently different to representational games, such as *Mario*, that have a diegetic world. As this work is focused on a third-person representational game with a diegetic world to explore, examples involving abstract games with little to no diegesis, such as *Tetris,* will be disregarded for now. Interactive video-game audio falls under five main headings: 'user-interface', 'dialogue', 'music', 'sound effects (SFX)', and 'ambience' (Taylor, 2010; Stevens & Raybould, 2016). User-interface refers to sounds generated from interacting with menus and any interface that is not within the scene; dialogue refers to any speech from the player, non-player characters (NPC), or narration; SFX includes sounds found within the scene, such as Foley, and external to the scene, such as notifications; and ambience refers to any background noise in a scene, such as room-tones for inside, or bird-calls and wind for a forest scene. Each sound found under these headings has a different role to play in the games' soundscape (Cheng, 2014), and further categorisation can help both sound and game designers understand the type of sound required for a given moment, and how, when, and why it should be used and for what purpose.

Audio in traditional media has often been categorised into two dimensions: 'diegetic' and 'non-diegetic' audio (Huiberts, 2010; Jorgensen, 2011). Diegetic sounds derive directly from actions or events within the scene (Curtiss, 1992; Chion, 1994), whilst non-diegetic refers to sound that has no primary relation to the narrative and no observable source within the scene (Gorbman, 1987; Tagg, 2004). In the context of video-games, examples such as the musical score or interface sounds from menus would be non-diegetic as they exist outside of the game-world, whilst footsteps, ambient sound, gunfire, voice radio messages, and certain types of notifications would be diegetic. However, as this terminology is derived from film theory, it is also built on the premise that, in film, a border exists between the story world and the real world (Grimshaw, 2007). With no such border existing in games, difficulty arises when applying the terms diegetic and non-diegetic to video-game audio, as there are examples that do not neatly fit into either definition. For sounds that fit neither construct, terms such as isomorphic (Curtiss, 1992), telediegesis (Grimshaw, 2007), and transdiegetic (Jorgensen, 2008), have been used. Prominent examples can be found in cartoons and animations, such as ascending/descending scales mimicking corresponding visual patterns, or cymbal crashes and musical stabs on impacts (Curtiss, 1992). This blend of diegetic and non-diegetic provides an additional layer of meaning in the context of video-games, and is especially significant when the participatory role of the player is taken into account (Collins, 2008; Jorgensen, 2011). The categorisations of video-game audio will now be explored, and how the terms diegetic and non-diegetic have been developed further in the specific context of ludomusicology.

Interactive video-game audio has previously been categorised according to its function and the information it conveys, how it is perceived relative to the game-world, or sometimes a combination of these approaches. Ekman (2005) explores the relationship between sound signs and signifiers in video-games as they relate to the player and the game-world. He discusses the relationships between sound signals, the sound itself, its denotative value, and sound referents, what the sound points to, its connotative value. Ekman expands the diegetic / non-diegetic classifications and divides video-game sound into four categories: ‘diegetic sound’, ‘masking sound’, ‘symbolic sound’, and ‘non-diegetic’. Diegetic sound refers to a sound, and its sonic reference, that exists in the game-world; masking sound refers to a sound that belongs to the game-world but contains a non-diegetic message; symbolic sound refers to the playing of a sound that is triggered by game mechanics but the sound itself does not belong in the game-world; and non-diegetic refers to sounds that do not belong to the game-world nor make reference to anything within the game-world. This categorisation retains the dichotomy of diegetic and non-diegetic, in that sounds are either present in the scene or they are not, but goes further to explain the function of sound as communication, and whether the information being communicated is to be understood internally or externally. Sound as a means of communication has been explored by Jorgenson (2006), who devises five main functions of interactive video-game sound: action oriented, atmospheric, orienting, control-related, and identifying functions. These add functional and experiential qualities to the game, by conveying information through proactive sounds, and supporting narrative through reactive sounds (Jorgensen, 2008). Collins (2007) further expands on the original dichotomy of diegetic and non-diegetic video-game audio, and proposes classifications relating to non-linear aspects of video-games. She argues that the terms commonly used to describe video-game audio, interactive, adaptive, or dynamic, are too vague, and that distinctions within these classifications are necessary. As such, she proposes the following categorisations:

Category	Description
Non-dynamic diegetic audio	The sound event occurs within the character's space, but with which the character has no direct participation.
Adaptive diegetic audio	Audio that adapts to the state of the game, such as different animals during night or day.
Interactive diegetic audio	The sound event occurs within the character's space, and the character can interact directly.
Non-dynamic linear sounds and music	The sound events occur outside the character's space and the player has no control over the playback. E.g. cut-scene.
Adaptive non-diegetic sounds	Sound events outside the diegesis that occur in reaction to gameplay but are unaffected by the player.
Interactive non-diegetic sounds	Sound events outside the diegesis that can react to players directly.

Table 2: Classifications of interactive game audio (Collins, 2007)

One of the most detailed models of video-game audio was designed by Huiberts (2010), who designed the IEZA (Interface, effect, zone, affect) model of video-game audio classification, see Fig 4. This theory maps four categories of sound in the game-world on two axes: ‘diegetic’ to ‘non-diegetic’, and ‘setting’ to ‘activity’. ‘Interface’ refers to the audio heard when a player interacts with a game medium such as a menu, and ‘affect’ refers to sound that communicates the game environment through social, cultural, and emotional references. These headings both describe non-diegetic audio, that is, sound that is outside of the fictional world that can only be perceived and interpreted by the player. ‘Effect’ describes sounds that are played in response to an in-game event or action, whilst ‘zone’ describes audio that is related to the visual setting of the game. The diegetic sounds found in the effect and zone domains can be conceptualised as audio that can be perceived by the avatar in the game-world. Examples include: footsteps, breathing, dialogue, weapon sounds, and collisions. These zones are not rigid, and sounds can move between them or have a function that is a combination of two zones.

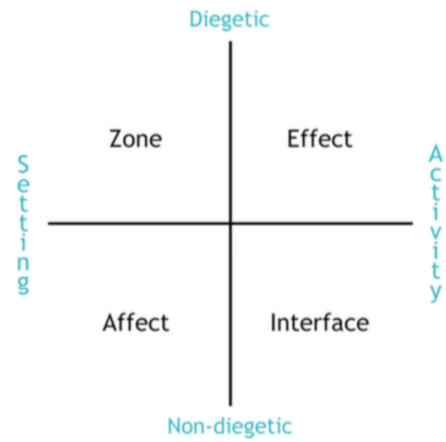

Figure 4: A representation of the IEZA model (Huiberts, 2010)

Many categorisations of video-game sound have been explored in the past, and though some bare conceptual similarities and definitional overlaps, the classification of sounds can help to highlight similarities, contrasts, and patterns (Schafer, 1994). In the context of this work, the definition of diegetic sound provided by Ekman (2005) will be adopted. Further exploration of these categories, and how the sounds function within them, can lead to a greater understanding of how sounds can be immersive.

2.2.3 Sound Functionality

Diegetic sounds function from both a usability perspective and an emotional perspective. Huiberts (2010) describes the usability of audio as optimisation, simulating what you would hear if you were actually in the game and providing necessary gameplay information to the player to assist their understanding of the game-world and ability to perform effectively within it (Marks, 2009). Optimising audio helps to create a believable and consistent game-world that the player can explore, providing aural feedback about the environment in place of physical sensations, and creating perceptions of spaces, locations, and times in the game (Grimshaw, 2007, 2012). Ambience, in particular, helps to provide a connection to how we perceive the real world, whilst also serving to block out the real world and prevent the player from being distracted (Marks, 2009). From an emotional perspective, or dynamism (Huiberts, 2010), diegetic audio generates emotional responses through atmosphere by building on the culturally established codes of Western media that are adopted through the viewing of film and television, and previous gaming experience (Roux-Girard, 2014). The importance of diegetic sounds, specifically their perceived reality within the game-world, and ability to signify real-time events, lifts them from the category of sonic decoration to being an integral part of the game's structure (Ekman, 2005). Footsteps, for example, are diegetic, and have a crucial role in playability, specifically helping the player feel present in the game-world; the texture of the surface beneath them is transmitted aurally in substitute of a physical feeling (Boyd, 2003).

The importance of diegetic audio as means of conveying information was explored by Jorgenson (2008), who examined what would happen if players were forced to play a game where all audio had been removed. The study involved 13 gamers who were asked to play one of two possible games, a real-time strategy game or a stealth-based game. After 20 minutes of gameplay, and unbeknownst to the player, audio was turned off, and the final 15 minutes of the experiment had to be played with no audio. Consequences arose for the game as both a virtual world and a usability system, with players reporting they felt like they were no longer in control, and that they could no longer orient themselves effectively in the game-world. Furthermore, from an atmospheric perspective, the mood of the current scene, and feelings of environmental realism, were all diminished, meaning progression through the game was slowed greatly. The work showed the importance of audio as a means of conveying information whilst the eyes are busy with other tasks, and also transferring sonic information from off-screen areas. Usher (2012) conducted a similar study. 12 participants were invited to play three different styles of game on a six-meter rear-projection screen with a 7.1 surround-sound system and adjustable lighting to provide a great gaming experience. Physiological measurements, heart rate, perspiration rate, and respiration rate, were taken throughout the experiment, whilst one participant would play through the three games with sound on, and the next would play with the sound off. Results showed that all participants that played with sound on had higher rates of physiological stimulation than those who played without. This was especially true for horror-style games.

Diegetic audio serves many functions, from: transmitting information, encouraging a suspension of disbelief, spatialisation and localisation in the environment, providing an awareness of the surroundings that are not visible through the screen, directing player attention, and rewarding the player (Marks, 2009; Stevens & Raybould, 2016). Now that the categories and functions of sound have been explored, it is possible to see how these relate to the general theories of immersion explored earlier and how sound can encourage immersion.

2.2.4 Sound as Immersion

As previously explored, immersion can be influenced by game components such as graphical quality and narrative. However, 98% of respondents in Huiberts (2010) survey on immersion confirmed that video-game audio has enhanced immersion for them in the past. During interviews with video-game sound designers, composers, and Foley artists, the term immersion is used frequently (Woelfer, 2015; TenFortySeven, 2016; FKN Ensemble, 2016). Many references are to the ambient sounds of the game-world, particularly in open-world games such as Uncharted 4: A Thief's End (*Soundworks Collection*, 2016a), Far Cry: Primal (*Soundworks Collection*, 2016b), The Last of Us

(*Soundworks Collection*, 2016c), the Halo series (*Soundworks Collection*, 2016e), the Assassins Creed series (*Steinberg*, 2016), and the DeadSpace series (*The Creators Project*, 2013). It is suggested that immersion in video-games relates to the level of sonic detail in an environment, and that the sounds most likely to induce immersion are the ones that the player does not consciously attend to, such as ambience, foliage, and background sounds, and how they are physically projected into the virtual world (Jonathan Lanier in Soundworks Collection, 2016a; Steve Papoutsis in The Creators Project, 2013). This lends weight to the argument that Foley sounds and ambience are processed subconsciously (Sonnenschein, 2001; Marks, 2009), and builds on the ideas of psychological immersion (McMahan 2003) and sensory immersion (Ermi & Mayra, 2005). Both of these terms describe the relationship between visual and auditory aspects of the game, resulting in a consistent and believable game-world that engages the players senses as though they were present in the game-world. However, audio can only be considered believable if it is properly synchronised with the visual world, referred to as 'synchresis' in film theory (Chion, 1994), and provides immediate and relevant sonic reactions to player-world interaction (Grimshaw, 2007; Guerraz & Lemordant, 2011); sounds should be indexical to the player's actions within the game environment as opposed to indexical to reality (Grimshaw, 2012). Immersive authenticity is achieved through the incorporation of detailed, representational audio, which helps to build a virtual reality that is based on the physical world, but contains references to cinematic and ludic conventions (Stevens & Raybould, 2015). Now that the functions of sound have been considered, it is important to consider how sounds can be designed and implemented into a game to achieve their maximum potential.

2.2.5 Considerations for the Sound Designer

Creating an immersive soundscape requires various technical considerations on the part of the sound designer, mixer, and audio programmer (Deutsch, 2011), and prior to audio middleware programs, there was limited access to any form of sophisticated mixing techniques in video-games (Bridgett, 2008). Sonic fatigue is one barrier to immersion, and results from excessive loudness (Somberg, 2016), unneeded repetition of audio that causes boredom (Sanger, 2004; Jorgensen, 2011), and unclear mixes that force the player to strain to discern individual sounds (Taylor, 2010). To avoid these issues, considerations such as equalisation (EQ), panning, and volume adjustments, can help to give each sound its own space in the mix (Patterson & Conway, 2014). Audio middleware programs, such as 'Wwise', can offer various solutions post-development (Somberg, 2016), including: sound randomisation at playback time to avoid repetition (Bernstein, 1997; Boyd, 2003), a combination of both active and passive mixing including mixer snapshots, HDR audio templates, on-the-fly EQ, voice-culling and sound prioritisation (Taylor, 2010), real-time, interactive DSP (Collins, 2008), and fully spatialised audio including 5.1 and 7.1configurations (Huiberts, 2010; Bridgett, 2008). Appropriate use of spatialisation and reverberation can provide sonic cues to help the player understand the type of space where a sound is occurring, and create a realistic and believable representation of that space; as a consequence, the player feels more immersed in that space (O'Keeffe, 2011; Patterson & Conway, 2014).

2.2.6 Summary

Sound as immersion has been explored by Jorgensen (2008) and Usher (2012), using two different methods that focused on playing games with or without sounds. Though these methods revealed the positive influence of audio on immersion, the binary nature of the experiments could be better refined to explore the influence of individual sounds or individual groups of sounds. Audio has many functions within a video-game, though classifying audio is sometimes problematic. The meaning of diegetic audio, in the context of video-games, is sound that is present within the game-world and that serves as a method of aural feedback, with each in-game action having an appropriate sound reaction (Marks, 2009). Unlike in film, where sound is left open for interpretation by the audience, aurally transmitted information provides relevant data for the player about how to interact with the game-world (Jorgensen, 2008). Foley and ambient sounds lend an air of believability to the game-world, and help to maintain and fulfil the expectations of the player (Collins, 2008). Many modern game engines present sounds that have depth, height, weight, spread, and are affected by physics such as occlusion, obstruction, Doppler, and the acoustic spaces of the game environment (Fay, Selfon, & Fay, 2004; Stevens & Raybould, 2016). The technical considerations outlined above combine with detailed and believable sound design to create an soundscape within the game-world that is increasingly detailed, consistent, and believable, and that actively engages the listener in the ongoing processing of aural information (Guerraz & Lemordant, 2008), ultimately resulting in a more extensive feeling of immersion.

2.3 Conclusions Drawn from Literature

To conclude the review of the literature, several key points have emerged that are especially relevant to this study, from the perspective of immersion as a general concept, game and sound design, and also methodological design. The literature has highlighted the multifaceted nature of immersion within video-games, and the various ways in which games can be considered immersive. On its broadest level, immersion in a video-game can be categorised into two dimensions: psychological immersion in the game environment, and perceptual immersion in the user environment (McMahan, 2003; Stockburger, 2003). Unpacking these further, perceptual immersion requires the user environment to be free of distractions, and to feature technology that can encourage immersion, such as headphones, surround sound systems, large screens, or projections (Usher, 2012; Brown & Cairns, 2004; Huiberts, 2010). Immersion in the game environment, however, requires a more focused examination, and has been categorised extensively in the past. The terms sensory, challenge-based, and imaginative immersion (Ermi & Mayra, 2005) have been adopted for this thesis to subcategorise psychological immersion. Multiple modalities of game design, including visuals, audio, narrative, character design, and controls, all work together to influence immersion in each of these categories (Cheng & Cairns, 2005). Though this thesis focuses on the role of sound, each of the modalities outlined above must be considered when designing the game to be as immersive as possible. This means the effects of modifying diegetic audio can be observed, else the player may be distracted by other elements of poor game design or the test environment itself.

Literature has highlighted the different types of diegetic audio, and the key roles that it can play in each of these immersive categories; creating detailed, sensory environments, providing auditory responses to actions and events, and developing mood and empathy through atmosphere. Both in the design and execution of this experiment, the psychological and perceptual stages of immersion must be considered. The game designed for this test must contain certain features to maximise the chances of psychological immersion, including:

- A basic narrative or plot (Douglas & Hargadon, 2000; Ryan, 2001; Martinez, 2014).
- Natural and fluid controls (Brown & Cairns, 2004; McGloin, Farrar, & Krcmar, 2013).
- An identifiable character (Ermi & Mayra, 2005; Grimshaw, 2012).
- An engaging and responsive virtual environment (Witmer & Singer, 1998).

To encourage immersion from an auditory perspective, the game must:

- Feature dynamic ambience that is location specific (Stevens & Raybould, 2015, 2016).
- Provide appropriate sonic responses to events and actions (Stockburger, 2003).
- Create realism using spatial representation, positioning, and reverb (Grimshaw, 2012).
- Synchronise sound and image (Chion, 1994; Guerraz & Lemordant, 2011).
- Develop atmosphere and mood (Brown & Cairns, 2004).

The experiment must create ideal testing conditions for perceptual immersion:

- The test environment must be comfortable and non-intrusive (Brown & Cairns, 2004).
- The participant must be alone in the room to reduce pressure (Cairns & Brown, 2005).
- The experiment must use technology that encourages immersion (Usher, 2012).
- The personality of the participant must be taken into consideration (Witmer & Singer, 1998; Wissmath, Weibel, & Groner, 2009).

Identifying the features of game design outlined above has allowed me to develop a game and test that will maximise the chances of player immersion, and the review has been crucial in informing my main research question and the development of the following methodology.

3. Research Methodology

Building on the work of previous studies (Brown & Cairns, 2004; Lazzaro, 2004; Cheng & Cairns, 2005; Jennet et al., 2008; Usher, 2012; Kirschner & Williams, 2014), it is possible to measure and quantify immersive gaming experiences under experimental conditions. Developing a bespoke testing environment can improve this process, as it gives experimenters the freedom to control almost every aspect of the game and test environment (Usher, 2012). This work used a three-phase convergent parallel mixed-method approach (CPMM) to understand the role of audio in creating immersion. The goal was to see if removing, modifying, or de-synchronising diegetic audio would be noticed by players, and if this would reduce their immersion in the game.

3.1 Research Strategy

This work used a CPMM approach, where concurrently collected quantitative and qualitative data is merged to provide a comprehensive analysis of the research problem (Creswell, 2014). The research had three phases:

1. To determine the candidates' likelihood of becoming immersed in the test game.
2. To record and observe live gameplay and record notes about questionable moments.
3. To view recorded footage, with the participant, in a semi-structured interview for discussion.

Phase one was achieved using a questionnaire that combined features of the Immersive Tendencies Questionnaire (ITQ) (Wimer & Singer, 1998) and Game Engagement Questionnaire (GEQ) (Brockmyer, et al., 2009), to gather quantitative data about aspects of gamer preference, genre expectations, and calculate the probability of becoming immersed during gameplay.

Figure 5: A participant as shown halfway through gameplay

Phase two involved ten participants, each of which was invited to play the game through a MacBook Pro connected to a 150-inch wall-projection whilst wearing ATH-M50X headphones, see Fig 5. The functions of the keyboard and trackpad were explained to those unfamiliar with a Mac, before the observer set the screen-capture software and video cameras recording and left the room. The screen-capture functionality of QuickTime software was used to record gameplay, as well as a GoPro Hero3+ Black Edition that was positioned to record both the player and the projection. The Wi-Fi functionality of the GoPro was used in conjunction with the app available on Android phones to monitor the live recording from a different room. This allowed the researcher to be removed from the room whilst still being able to monitor the experiment and make notes of points in the video that could be re-examined during the interview. Finally, phase three involved a semi-structured, open-ended interview to gather qualitative data from the player's perspective (Creswell, 2014) about their views of the game, the test environment, and, most importantly, how they felt the changes in audio affected their immersion in the game, or if they even noticed the changes at all. Using a CPMM approach allowed me to draw on the strengths of both quantitative and qualitative research and reduce limitations a single method may incur, providing a deeper and more complete understanding of the research problem (Creswell, 2014). This was ideal for research into audio-based immersion as the field is relatively new, and it allowed me to gain insight about any problems from the view of the participant. However, the validity and accuracy of the method is compromised somewhat by its requirement for extensive data collection, the time-intensive nature of analysing two sets of data, the subjectivity and interpretation of qualitative data imposed by the researcher (Wengraf, 2001), and the large sample size required (Creswell, 2014). This work hopes to combine previous methodologies and incorporate data triangulation (Bell & Waters, 2014) to test assumptions and validity.

3.2 Analytic Tools

The three phases of the experiment required different analytic tools. The following passage will explore each of these in turn, discussing previous experiments that have used them and the benefit and shortcomings of using them for my work. It was important to understand the participants' susceptibility to becoming immersed, as well as their preferences towards gaming (Jennett, et al., 2008). This helped to determine if any breaks in their immersion during gameplay were due to experimental conditions, or if they were due to certain personality traits (Wissmath, Weibel, & Groner, 2009), a disinterest in the genre of game (Brown & Cairns, 2004), an inability to connect using an unfamiliar controller (McGloin, Farrar, & Krcmar, 2013), or any other external or internal factors likely to cause distraction (Ermi & Mayra, 2005). Two questionnaires proved useful for determining this: The Immersive Tendencies Questionnaire (ITQ), designed by Witmer and Singer (1998), and the Game Engagement Questionnaire (GEQ) designed by Brockmyer et al. (2009). Building on these works, a new game engagement questionnaire was developed, see appendix 3, to measure elements of both immersive tendencies and gamer preference. This questionnaire revealed which participants would be most susceptible to entering immersive states, and also nullify any data resulting from breaks in immersion that occurred as a result of game features other than the ones I hoped to observe. The inclusion of the term immersion was avoided, where possible, so as to not cause confusion about its meaning (Ijsselsteijn, et al., 2000).

Video recording has many benefits in the test environment, especially in regard to observational research regarding behaviour, content, processes and interactions (Wu, Guo, & Seaman, 2009; Bell & Waters, 2014). However, observational methods can be risky due to the perception and interpretation of what is seen or said, and certain precautions must be taken to avoid introducing bias through “preconceived ideas and prejudices” (Bell & Waters, 2014, p.211). Research has shown few significant benefits from video-recording in relation to immersion and video-games (Cheng & Cairns, 2005). As such, footage recorded in my experiment was not analysed from the perspective of body language and facial expression, though it has become a more established practice in video-game research (Lazzaro, 2004). It was, however, used to raise questions and implications during the interview in phase three.

Using interviews to extract information about subjective experiential phenomena is becoming a more widely accepted practice (Bell & Waters, 2014). In regard to immersion, interviews have been used to build an insight into negative gaming experiences through dialogue with the players (Jorgensen, 2008; Jennet, et al., 2008), and to reveal insight into gaming experiences that have been ruined by inappropriate audio (Huiberts, 2010). In some cases, interviews have proved more reliable than analysing the psychophysiological variables, such as heart and perspiration rate, that are associated with immersion (Nacke, Grimshaw, & Lindley, 2010). In these examples, the interviews were always conducted shortly after the initial experiment or gameplay session took place. This meant players could easily recall aspects of their gameplay experience, and reduced the chances of participants feeling discomfort through an inability to remember certain events in great detail (Bell & Waters, 2014). It also has the added benefit of creating fresh memories in the player's mind to discuss, instead of retrieving previous memories of gameplay that may be less vivid, and helping them to relax (Brown & Cairns, 2004). Using both observational and interview techniques allowed me to gain a level of insight into video-game immersion that cannot be rivalled by quantitative means alone (Bell & Waters, 2014). Conducting the phases in the order I chose meant that the player could be left to play the game uninterrupted for the length of the experiment. This was a crucial decision, as it avoids breaking the immersive state of the player to ask questions (Cowley, et al., 2008), and allows the researcher to be absent from the room, avoiding unnecessary pressure on the player that may make them perform differently (Creswell, 2014).

3.3 The Game

In order to have full control over game variables that are being explored, a new game-world was created using the Unity game construction engine and Monodevelop scripting environment. My main goal was to create a game that looked, sounded, and played like any other game, and not one that had just been designed for an experiment. This meant that it needed a start menu, clear instructions when starting a new game, and a storyline to follow; controls needed to be easy to pick up, the visuals and audio had to be engaging and dynamic, and there needed to be high levels of user-interaction within the game environment. For inspiration, I looked to games such as *Jade Cocoon* (Genki, 2000), *Jade Cocoon 2* (Genki, 2004), and the *Final Fantasy* series *(*Square, 1987). The game was given a brief narrative and back story, to provide some context for the player. As a knight sent on a quest by the King, the objective of the game was to navigate through the environment, collecting treasure, items, and solving basic logic puzzles along the way, with the ultimate goal of passing through a teleportation device, whilst avoiding enemies, to move to the next area and find more treasure to take home. The game was deliberately designed to be very visually and sonically engaging, and provide lots of opportunity for user-environment interaction so the player could become involved in the game mechanics. These design choices provided a large enough variety of stimuli to increase the chances of becoming immersed (Nacke & Lindley, 2008).

When choosing the audio to be modified, I needed sounds that could be repeatedly altered at various times, without knowing what the player would do or where they were. As such, the choice of sounds needed to be omnipresent throughout the gaming experience, and inherently linked to the character so that opportunities for modifications could arise at any time – footsteps, therefore, seemed the most appropriate choice. It was suggested that any changes made must air on the side of discretion, otherwise any break in immersion may be down to inherent bizarreness as opposed to what the study is focusing on (Cheng & Cairns, 2005). This view guided my choice in modifying footstep sounds and ambience, as they are already a discrete, albeit important, sonic element of the game-world.

When designing the sounds for the game, I focused on encouraging a sense of immersion through detailed sonic environments (Jorgensen, 2011). Ambient sounds were static loops recorded in appropriate locations, see Fig 6. These loops were then reduced to an appropriate file size and length, and combined with one-shot sounds such as bird calls, water droplets, and wind howls, that were triggered randomly and also randomly positioned in three-dimensional space to create a dynamic and constantly changing virtual soundscape (Stevens & Raybould, 2016). Footsteps covered two surface types, and featured five samples for each foot per surface type, the playback of which was randomised when called, resulting in 20 footstep samples used and 50 possible playback combinations.

Figure 6: Recording ambience for the forest environments

Sound parameters such as occlusion, obstruction, volume, and frequency shelves are modified in real-time to closely reflect how sounds would be heard in real life relative to distance. As the player approaches a fire, for example, intimate sounds of whistling and crackling are triggered within a certain distance, but are lost as the player gets further away. A higher level of detail in the ambient sounds gives the player a more complete sense of being spatially located in the game-world, and matches the auditory expectations of the player to the conventions of the game-world (Grimshaw, 2012). All audio was recorded using a Zoom H6, equipped with a mid-side or X/Y microphone attachment, at 16bit 48kHz for audio fidelity and the option of sample-rate reduction should it be required. Samples were then trimmed to the appropriate size and cleaned, removing extraneous noise, hisses, crackles, and pops, using Logic Pro X, before being implemented into the game using the audio middleware program Wwise. During play, as the player moved the character through set trigger points in the map, footstep sounds would be removed, modified (by filtering until unrecognisable as a footstep), or de-synchronised for a period of time, and then returned to normal. Ambient audio would also be discontinuous in certain places; the sounds of different environments would blend together instead of cleanly changing from one to another. The aim of this was to trigger a break in immersion through the unnaturalness of the game environment and lack of avatar-related sounds (Stevens & Raybould, 2016).

Figure 7: The natural path of progression through the level

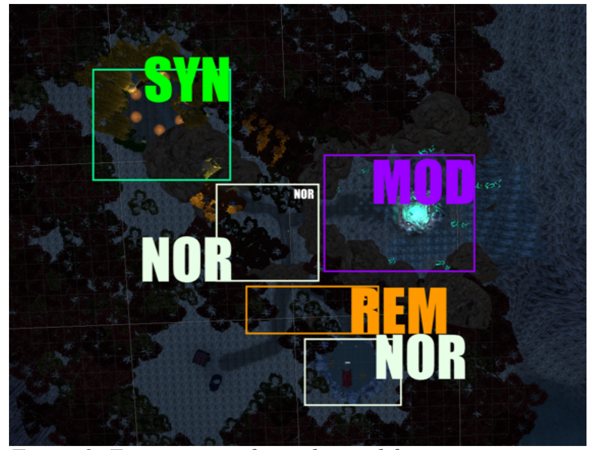

Figure 8: Trigger zones for audio modification

Figures 7 and 8 show birds-eye views of the game map. Trigger zones were placed along the natural path of game progression the player would have to follow, shown in Fig. 7 by coloured arrows, in order to complete the level (Gard, 2010). The trigger zones, which are invisible to the player, are shown in Fig. 8, and are labelled according to their function:

NOR = Footstep audio normal / returned to

REM = Footstep audio removed

MOD = Footstep audio modified / discontinuous ambience

SYN = Footstep audio de-synchronised

A playable copy of the game has been included with this thesis, though it is only compatible with Mac. A video example from the study has also been included, with reference notes to the times in the video that audio is modified. To view both these, and all the sound files used for the project, please see the accompanying disk.

3.4 Participants

Ten people participated in this study. They were selected because they regularly played video-games for more than 10 hours per week on average. The sample included nine males and one female, with an age range between 21-27, all of whom currently owned a games console or a PC. Though distribution of the consoles was not even; six participants played games on an Xbox console, two on a PlayStation, and one on PC. Participants were asked to sign a consent form, shown in Appendix 3, agreeing to take part in the experiment and confirming their understanding of what the experiment would consist of, as well as their acceptance of the images that would be projected on screen and the video-recording of themselves.

A pilot test was conducted on two gamers, who were not included in the final study to avoid bias (Bell & Waters, 2014), to determine if the data returned from the questionnaires was usable and if the game had any issues which raised serious concern in terms of research focus.

4. Results

4.1 Analysis

Phase one provided quantitative data from a questionnaire. All participants reported becoming deeply involved in narrative mediums to some degree, though only two reported experiencing deep-involvement frequently whilst the remaining eight experienced it occasionally. Only one participant reported frequently becoming involved to such a degree that it becomes difficult to get their attention, whilst six reported it occasionally, and three had never experienced it to their knowledge. Every participant reported frequently losing track of time whilst playing games, though the genre and style of game varied from person to person. The games reported included: sport, third-person action and shooters, open-world exploration, free-roam, fighting simulation, and adventure. Empathy towards a game character was evenly distributed; four reported never experiencing it, four occasionally, and two frequently, though the respondents who did report emotional involvement reported it happening most frequently with combat and action games.

Responses to questions were split into subscales based on the work of Witmer and Singer (1998): the tendency to become involved in activities (INVOL), the tendency to maintain focus on current activities (FOCUS), the tendency to play video-games (GAMES), and, an extra category for the purpose of this work, the technology normally used to play games (INPUT). An average score for each subscale was calculated and presented as a value out of five. The total score from all subscales was presented as a percentile probability of becoming immersed in the current game, considering the personality traits of the individual, the game type, controller type, and playing environment. While most questions provided ordinal data, questions Q2, Q3, Q11, Q13, and Q14 provided categorical data relating to genre preference and general gaming habits. These questions were analysed differently, though they still contributed to the overall score for each subscale due to their relevance. For example, Q2 asked about the preferred platform for playing games, and Q3 asked about preferred controller. Where the responses to these questions did not match the experiment variables, a value of 1 was used, and where the response did match the experimental variables, a value of 5 was used. Q11, Q13, and Q14 also provided categorical data, relating to each participant's favourite games, and games in which they have frequently experienced feelings of immersion. The Jaccard similarity index (Stephanie, 2016) was used to determine similarity coefficients between the test game and any games referred to by the participant in their responses. By comparing how many game variables were present in both the test game and the games referred to by players, taking into consideration variables such as graphics, mechanics, functions, objectives, gameplay style, environments, narrative, camera perspective, and level of interactivity, a percentage of similarity was calculated. This degree of similarity was then scaled between the range of 1 to 5, to return a variable that could then be used in the original average-based calculations. The average percentage of immersive likelihood was 61% over the group, indicating a slightly higher-than-average chance of immersion in the game.

Phase two provided audio-visual gameplay footage. With the main focus on creating timestamps and written notes of key moments during recording to raise in discussion with the participant during the interview, the material itself required no deeper analysis. The audio-visual material provided an extra layer of mental stimulation for the player, to generate more valid memory and points of discussion from the perspective of the participant. Very few key moments were highlighted during gameplay, apart from the recurrence of participants having to look away from the screen to the keyboard to find the right button to press. The main data gathered from the audio-visual material was the specific time in the video that players entered the trigger zones. Having these reference times enabled me to quickly skip to those points in the video during the interview, without wasting time going through unnecessary footage.

Phase three provided qualitative reports on the game playing experience from the perspective of the player. All participants reported enjoying the game, though nine commented that the computer keyboard was uncomfortable to play on. This correlates with answers given during the pre-play questionnaire relating to console and controller preference. All participants stated that the game was engaging, and that they became involved in the narrative and tasks, though none reported feeling emotionally responsive to the character. Participants noted a variety of game-elements that they felt were confusing or that could be improved, these will be explored further in the critical reflection. Importantly, after demonstrating the changes in audio, not one participant noticed any of the three changes made to the footstep audio, nor the discontinuity of the ambient audio.

4.2 Findings

As a result of immersion, participants failed to notice what had been hypothesised to be modal incoherence – a mismatch in the soundscape of the game-world compared to the expectations of the player that had been formed through previous exposure to games and reality (McMahan, 2003; Collins, 2008). This result extends the previous findings of Brown and Cairns (2004) and of Cheng and Cairns (2005), and reiterates that whilst modal incoherence could be a barrier to immersion, its importance is diminished once immersion has already been achieved. Even in regard to diegetic audio, an important element of game design that encourages sensory immersion and presence within the game-world (Stockburger, 2003; Grimshaw, 2012), players remained undeterred to changes that should, in theory, have diminished their belief of the game-world and its consistency (Guerraz & Lemordant, 2008).

When asked why the participant thought they did not notice the changes in audio, the most common answer was that they were too busy focusing on the task presented in the game:

<u>P2, Q9/10/11</u>

> I didn't notice, [sic] too busy working out how to light the bloody torch! When I'm given an objective, I focus on that objective. If there were no other nature sounds then it might have been more noticeable. I really liked the nature sounds at first when the game started, I'd never played with headphones before so I could hear everything. But then I had to focus on instructions and moving and doing stuff, maybe that's why I didn't notice the sounds disappear. Or I was too busy looking around, the game looks really nice, I don't know.

Though other answers included: focusing on reading the subtitles:

P4, Q9/10/11

> The instruction writing was appearing so fast that it was hard to read, I was focusing really hard on that because I knew once it disappeared I wouldn't have a clue what to do after.

In relation to the forest ambience, most players thought that it did not sound out of place in the cave setting, and that one would still be able to hear the birds outside from inside the cave:

P1, Q12

> I didn't notice it, but it's not that out of place, I think you'd still be able to hear the forest from inside, wouldn't you?

P2, Q12

> I didn't notice. Can't birds be in caves?

It appeared that the controller was of more concern to the participants, with nine stating it had a negative effect on their immersion in the game:

P8, Q3

> The keyboard was really annoying. It wasn't, like, the difficulty of the controls it was just finding a way to use them that was comfortable, I couldn't really get into it [the game] because I kept needing to look at where my fingers were going.

Discussing the de-synchronisation of footsteps, one participant noted that the speed of the character's movements may have been a contributing factor:

<u>P10, Q10</u>

> Now that you've pointed it out I can tell, but when I'm watching or playing the guys legs are moving so fast anyway it's hard to tell when his foot hits the ground apart from when he stops and you hear two or three more steps.

Whilst another noticed the lack of relevance to the current task:

<u>P7, Q10</u>

> I've never really paid attention to what their [characters] feet sound like. Other games, maybe, like, stealth ones, where you've got to sneak or listen for guards or something, but a game like this I wouldn't notice.

Returning to the research question, results showed that removing, modifying, or de-synchronising some diegetic elements of the soundscape in a third-person RPG does not have a negative influence on the player's immersion in the game. Participants believed this was due to the need to focus on tasks, and their discomfort with the controller. This opens many avenues for discussion, regarding the structure of the methodology, implications for sound designers and game design, and implications for future research into the field of ludomusicology.

4.3 Discussion

Contrary to the view of Marks (2009), who believed that any changes in diegetic audio would “stick out like a sore thumb”, the results showed that participants did not notice the changes and discontinuity in diegetic audio presented in this experiment. I offer several explanations and interpretations for this. Firstly, the footsteps had less relevance to the player due the character being non-realistic; the usability functions of the diegetic sounds were redundant due to the players attention being focused elsewhere; diegetic audio has reduced functionality in third-person games compared to first-person; being immersed from a challenge-based perspective can reduce sensory immersion; and, finally, a natural speed of progression through the game over a longer period of time would have revealed different results.

Given that the function of footstep sounds and ambient audio is to spatially locate the character in the game-world and provide a layer of conceptual realism (Marks, 2009), participants reported no issues with either of these modalities. The de-synchronisation of footsteps should have made the character feel less responsive to the directions given by the player, given that the tight synchronicity between visuals and sounds and the responsiveness of the environment lend to the believability of the virtual space (Chion, 1994; Grimshaw, 2012). One participant pointed out that the character and animations used for the character were heavily cartoon-based, and as such the leg movements of the character were quite rapid and the exact moment of footfall was not very clear. In this respect, the de-synchronisation of footstep sounds may have had more of a negative impact in a game that features characters more closely related to real life, building on the ideas of authentic and inauthentic sounds (Stevens & Raybould, 2015). Within the specific context of third-person video-games, the usability functions of diegetic audio in transmitting information may be of less relevance than in first-person perspective games. In first-person games, that typically provide a viewing angle of between 60-70 degrees (Stevens & Raybould, 2015), auditory cues can provide important information about surroundings that are situated outside of the viewing area (Fay, Selfon, & Fay, 2004). In the third-person perspective, however, the player can already see much of the virtual space around the character, and as such the need for audio cues may be reduced (Taylor, 2002). Furthermore, discussion with participants revealed that players do not consciously listen for sounds unless the game dictates they do so, or they have a purpose for doing so, if they were in danger or if it was a requirement of a task, for example. This builds on the theories of sound functionality, usability, and location in virtual space, as well as the perceptual systems of hearing. The usability of diegetic audio is normally found in reference to predicting danger, much like in real life (Huron, 2014), and gathering information about the virtual world. But once the acoustic conventions of the game-world have been processed, it appears they do not need to be consciously re-processed, as the mental model of the environment has already been formed. Having found that little attention is directed towards the diegetic audio in a third-person game, unless it is providing a function within the game, it is clear that the attentional direction of the player contributes to the type of immersion they experience in the game-world.

The tutorial-style level used for this experiment allowed the player to become familiar with game controls, mechanics, visuals, and auditory styles. This process requires greater cognitive focus than when the player has already learned the controls and conventions of the game-world and is playing a level later in the game (Suddaby, 2012). Tutorial levels can provoke challenge-based immersion, or flow, through the act of learning and mastering controls, and responding to instructions that are appearing on screen (Czsentmihalyi, 1991). The need to focus consciously on the keyboard, the placement of fingers, whilst simultaneously reading instructions from the screen, reduces the attention available to give to other game elements (Calleja, 2011; Lemarchand, 2012). This confirms the view that if a player is immersed from a challenge perspective the likelihood of acknowledging shortcomings in feature that promote sensory immersion is reduced (Westerberg & Schoenau-Fog, 2015), and that the presence of inauthentic audio is only accepted when the drivers for immersion are challenge-based, and the sounds have a clearly defined ludic role (Stevens & Raybould, 2015). Furthermore, although most of the participants quoted the controller as being one of the main distractions from the game, they were sufficiently engaged to be oblivious to the test, suggesting that a low level of immersion, even engagement or engrossment, is sufficient to overcome incoherent or discontinuous diegetic audio to a certain extent (Brown & Cairns ,2004; Cheng & Cairns ,2005). This also supports the view that sensory immersion, of which diegetic audio plays a key role, only has a supporting role during gameplay, one that merely contributes to both challenge-based and imaginative immersion (Ermi & Mayra, 2005; Huiberts, 2010).

Examining the work of Huiberts (2010), the PUGS questionnaire reveals many instances of sounds that ruined players' immersive experiences of games. Compared to the results from this thesis, the sample size needed to be larger in order to provide conclusive data. However, it could also relate to the amount of time and effort that the player has devoted to the game. Respondents to the PUGS questionnaire spoke of games that they had already become deeply familiar with, and had already spent time and energy playing. Their acknowledgement of inauthentic sound, therefore, could have been built gradually over time and with repeated exposure to the sounds in a variety of game environments and situations. As players progress naturally through a game they learn about its design features and deepen their understanding and perception of, not only the game that they are playing, but game features in general (Kirschner & Williams, 2014). This natural progression through the game takes time, but suggests that, if given longer to play during the experiment, the participants of this study may have adapted to game controls, resulting in a more fluid gaming experience, and allowing players more opportunities to become emotionally connected with the character and game environment. Furthermore, increased play time would provide more opportunities for exploration that is not objective-focused, freeing the players' attention to wander and explore the game environment, possibly highlighting the issues that were presented by this study.

The findings bring forward several arguments and implications for future work in the field of video-game research, and the creative design process for sound and game designers. For sound designers, one of the most important stages of implementing audio into a game is the mixing stage (Taylor, 2010). This is the point in the development of the game where each sonic element intended to be present in the game is balanced against one another. During this phase, the sound designer and games designer work closely together to decide which sounds they want the player to hear at that moment in time and what functions each sound will serve when it is audible (Stevens & Raybould, 2016). Research such as this can help determine where the focus of a player is directed at any given moment, and if the mixing of the audio can be implemented differently knowing that there are some sounds the player will be less aware of due to their attention being diverted elsewhere. This would also apply to equalisation choices for the sound designer; if a sound is occupying a certain frequency range that could be better occupied by a sound of higher importance, then it can be cut depending on the focus of the player. Choices such as these can help in deciding which sounds should have priority in a given scene, and which sounds could be pushed to the back of the mix or cut completely. From the perspective of the game designer, it also helps in understanding the functions that different sounds have within the game-world and if these functions are what the game needs at that moment. Experiments such as this could also be used during the research and development phase of video-games, to ensure the creative decisions being made are going to return the best results. Finally, time-constraints on video-game projects are tight, and optimising workflow is important. On the part of the sound designer, research such as this can help prioritise their time, and inform how it is allocated between the development of different sounds.

For researchers investigating topics such as immersion in video-games or video-game audio, this research has shown the power and flexibility offered by designing a game-world specifically for an experiment. Using games that have been designed previously allows greater familiarity from the players perspective, and may overcome issues of controls, narrative, or character empathy, but the researcher is then limited by the number of game variables available to modify and explore. This is especially true when exploring video-game audio, with the only options being to present the sound as off or on (Jorgensen, 2008; Usher, 2012). Though the results from this work have highlighted many topics for discussion, a critical reflection on the development of both the game and the methodology will help to highlight any issues that could be addressed in future work.

4.4 Critical Reflection

On reflection, both the design of the game and the methodology presented several inconsistencies that need to be addressed, though they do highlight key points for future studies to build upon. The shortcomings in the design aspects of the game need to be acknowledged: balancing task-based and explorative environments, overcoming player-avatar empathy, using a character that suits the needs of the experiment as well as fits the design aesthetic of the game, and appropriate control input. Although the game was designed to provide a variety of tasks and situations to encourage immersion, on reflection, designing a tutorial-style level promoted challenge-based immersion, and thus reduced the opportunities for the player to become immersed in the sensory aspects of the game as this work hoped to observe. A future experiment would need to provide equal amounts of challenges and opportunities for solo exploration. All participants reported difficulty empathising with the character. One way of overcoming the lack of player empathy with third-person characters is to provide them with a chance to customise the look of their character; choose the clothes they wear and the weapons they own; the companion they quest with and so on. This would help to encourage more empathy and emotional connection, or imaginative immersion (Ermi & Mayra, 2005), between the player and character in a shorter time frame. Studies have shown that incorporating avatar customisation results in greater feelings of enjoyment (Teng, 2010; Frasca, 2001; Kim, et al., 2015; Schmierbach, Limperos, & Woolley, 2012; Smahel, Blinka, & Ledabyl, 2008). The style of the game was intended to be fun, engaging, and mystical, although the design of the character proved ineffective in monitoring the de-synchronisation of footsteps due to the speed at which the character moved. This could be rectified by using a more life-like avatar.

From a methodological perspective, mixed-method designs have proved especially useful in the investigation of video-games (Cheng & Cairns, 2005; Jennet, et al., 2008; Jorgensen, 2008). However, concerns have been raised about the validity of Likert scales for measuring subjective measures (Sullivan, 2013), and the bias that can be introduced from the researcher's perspective during subjective interviews (Bell & Waters, 2014). The questions presented in both the pre-play and post-play phases of the experiment could have been refined further, especially the wording of certain questions that caused confusion. In regard to the audio modifications presented by this work, more extreme modifications could have yielded different results. The footstep sounds, for example, could have been changed to a completely different sound sample, or, instead of reducing the volume, the volume could be increased. This would allow the threshold to be determined whereby players start to notice inconsistencies.

Finally, the results of this experiment can only relate to gameplay experiences that occur in the optimal environment designed for this study. It should be noted that in everyday gaming situations, many gamers do not have access to the audio-visual technology used in this experiment. Some may have, but the configuration of their setup may be different, with less focus on audio and more on the visual elements of the gaming system, such as listening through television speakers. Such a system would result in the auditory changes presented in this test being even less noticeable and, potentially, less damaging to the immersive state of the player. By discussing the findings and acknowledging the limitations and issues that arose throughout the course of this work, suggestions can be made that would be helpful in the design of future research that focuses on similar topical areas.

4.5 Future Research

This study had to be limited in its scope due to the constraints of time and the manpower required to see it come to fruition. Without such constraints, the topics and findings from this work could be explored further and in much greater detail, for example: the role of diegetic audio could be examined during moments when the player is not currently focused on a task, or when diegetic audio is required to complete an objective. In regard to audio-based immersion and the functions of certain diegetic sounds to provide information, a further question arises: Would players notice diegetic sounds being removed, modified, or de-synchronised in a game if the diegetic sounds were needed to complete a task? When considering the responses of participants, and the previous research in this area, when a sound has a specific function or value from a usability perspective, more attention is directed towards it. Future experiments could examine this theory by gradually increasing the level to which sounds are modified, or increasing the number of sounds that are removed from the game when aurally transmitted information is of most importance to the player. In this scenario, more considerations from a design perspective would be: increasing the length of time participants get to play, allowing players to bring their own controllers, and using a level from a game that had a more balanced variety of tasks and chances for exploration. The methods used in this work are not limited to third-person games, and could be used to examine first-person or virtual reality games, or even compare the functions of sounds between games that require different perspectives. Virtual reality arguably provides a more immersive experience than games presented on a screen (Patrick, et al., 2000; Ruddle, Payne, & Jones, 1998). With the recent rise of VR headsets and the promotion of VR games, research into the field of audio within virtual reality games could become especially relevant.

Though this study sought to observe the psychological effects of immersion, as opposed to physiological, a larger test could combine the two. Audio has been shown to contribute to immersion through physiological variables: higher arousal states, increased breathing, heart, respiration, and perspiration rates, were all noted as a result of playing a game with audio versus without (Jorgensen, 2008; Usher, 2012). The physiology of the eye can also indicate immersion, though this was not investigated from an auditory perspective (Jennet, et al., 2008). A combined research approach, from both a psychological and physiological perspective, would provide an even richer data set for analysis.

5. Conclusion

The intention of this work was to evaluate if removing, modifying, or de-synchronising diegetic elements of a game's soundscape would have a detrimental effect on a player's immersion in a third-person game. To investigate this, I developed a game-world that featured many of the immersive elements of gameplay explored in this work, such as: detailed environments, tasks and objectives, a rudimentary narrative, and high levels of user-environment interaction. Built into this game were trigger zones that would remove, modify, or de-synchronise elements of diegetic sounds as a player passed through them, in an effort to break their sensory immersion and believability of the game-world. Conducting an interview immediately after gameplay and entering into discussion with the players aimed to offer insight into the gaming experience from a players' perspective.

Though previous work had highlighted the inconsistencies in disrupting immersion

(Cheng & Cairns, 2005), it was hypothesised that, due to the importance of audio in creating believable game-worlds (Collins, 2008; Grimshaw, 2012), inconsistencies in diegetic audio would have a negative effect on immersion, specifically sensory immersion (Ermi & Mayra, 2005), and consequently reduce the enjoyment of the player. This research showed, however, that once immersion has been achieved, inconsistencies within the soundscape can go unnoticed if they are not directly related to the tasks or objectives of the player. All ten players that took part in this study failed to notice that footstep sounds had been removed, modified, or de-synchronised, even though the changes created obvious discontinuity in the game's soundscape and were hypothesised to reduce the sense of presence. Nor did players notice discontinuity in the ambient backgrounds of the world when two sonic environments blended into one, creating an incoherent environmental soundscape. Discussion with participants revealed that the objective-based nature of the game meant that all their attention was directed onto learning game controls, reading instructions, and working out how to complete tasks. This suggests that, while diegetic audio is an important part of creating a believable game-world and provoking sensory immersion, when the attention of the player is diverted by an objective, it becomes less relevant: that is, unless the diegetic sounds provide immediate feedback about a task or are directly related in a way that aids the player in their completion of said task. Ultimately, understanding the requirements of the video-game being designed is crucial. Whilst this work offers insight into sound design and creating immersive worlds, not all games are required to induce immersion, and nor is it required at all times during play. Though when it is required, the inclusion of immersive elements of gameplay can improve the gameplay experience for the players.

For sound and game designers, this work has shown how the focus of the player can make them oblivious to certain changes in the acoustic environment of a game. This raises implications regarding decisions during the mixing stage, design of sounds, and how the time of the game-sound designer could be best spent. For researchers, this work has shown that developing a game specifically for a research project allows more freedom over the choice of sounds that can be investigated and the ways they can be modified. It also enabled a move away from the binary examinations of audio that had been present in previous research, where sound was either on or off. Through the development of a specific game environment and use of deleterious usability, this research offers further insight into the role of diegetic audio in third-person perspective games, and the role that it plays in creating immersive gameplay experiences.

Appendices

Appendix 1 – Participant Information Sheet

Participant Information Sheet

Research Project Title:

Name of Researcher: Matthew Varley

Contact Details of Researcher: 07960865956 / u1263633@hud.ac.uk

You are being invited to take part in a research project. Before you decide, it is important for you to understand why this research is being done and what it will involve. Please take time to read the following information and discuss it with others if you wish. Ask if there is anything that is not clear or if you would like more information. May I take this opportunity to thank you for taking time to read this.

1. What is the purpose of the project?
The project is intended to provide the research focus for a module which forms part of my degree. The main purpose of my research is to examine player experience in a range of technological environments.

2. Why have I been chosen?
You have been chosen as you met the requirements ascertained from the online survey you completed; you are a frequent player, or developer, of video-games; you prefer the third-person RPG genre; you have frequently experienced a loss of time, or transfer of consciousness, whilst playing video-games.

3. Do I have to take part?
Participation on this study is entirely voluntary, so please do not feel obliged to take part. Refusal will involve no penalty whatsoever and you may withdraw from the study at any stage without giving an explanation to me.

4. What do I have to do?
You will be invited to take part in a short gameplay session at the University of Huddersfield. A video camera will be set up to record you playing, and screen capture software will also be used to record gameplay footage. Afterwards, you will be invited for a short interview whilst re-watching the footage. This should take no more than one hour of your time from arrival to departure.

5. Are there any disadvantages to taking part?
There should be no foreseeable disadvantages to your participation. If you are unhappy or have further questions at any stage in the process, please address your concerns initially to me if this is appropriate. Alternatively, please contact Professor M. Adkins (**m.adkins@hud.ac.uk**) at the School of Music, Humanities and Media, University of Huddersfield.

6. Will all my details be kept confidential?
All information which is collected will be strictly confidential and anonymised before the data is presented in any work, in compliance with the Data Protection Act and ethical research guidelines and principles.

7. What will happen to the results of the research study?
The results of this research will be written up in my MA thesis. If you would like a copy please contact me.

8. What happens to the data collected?
The data collected will be used in my thesis to support or disprove a hypothesis.

9. Will I be paid for participating in the research?
No payment will be made for participation; though refreshments and snacks will be available for participants before the experiment begins. If anyone has any dietary requirements, please let me know and arrangements can be made.

10. Where will the research be conducted?
Research will be conducted at:
The Creative Arts Building,
University of Huddersfield,
Huddersfield,
HD1 3DH

12. Who has reviewed and approved the study, and who can be contacted for further information?
Dr Elizabeth Dobson
Senior Lecturer in Music Technology,
School of Music, Humanities & Media,
University of Huddersfield,
Huddersfield,
HD1 3DH
Email: e.d.dobson@hud.ac.uk
Tel: +44(0)1484 471890

Appendix 2 – Participant Consent Form

Participant Consent Form

Title of Research Study: A Case Study on Understanding the Influence of Diegetic Audio on Immersion in a Third-Person Role-Playing Game

Name of Researcher: Matthew Varley

Participant Identifier Number:

- [ ] I confirm that I have read and understood the participant Information sheet related to this research, and have had the opportunity to ask questions.
- [ ] I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.
- [ ] I understand that all my responses will be anonymised.
- [ ] I give permission for members of the research team to have access to my anonymised responses.
- [ ] I agree to take part in the above study

Name of Participant:

Signature of Participant:

Date:

Name of Researcher: Matthew Varley

Signature of Researcher:

Date:

Appendix 3 – The Pre-Play Questionnaire

1. On average, how many hours per week do you play video-games?

   ____

(INPUT)

2. Do you have a preferred video-game platform?
   - ☐ PlayStation
   - ☐ Xbox
   - ☐ Nintendo Wii
   - ☐ PlayStation Vita
   - ☐ Nintendo DS
   - ☐ PC/Mac
   - ☐ Other, please specify:

   …………………………………………

(INPUT)

3. When playing games, which controller do you most often use?
   - ☐ Control pad
   - ☐ Joystick
   - ☐ Keyboard and mouse
   - ☐ Other, please specify:

   …………………………………………

(INPUT)

4. Do you have experience of playing games on a computer using a keyboard and mouse?

   Never 1 → 5 Frequently ___________

(INPUT)

5. Do you prefer playing video-games whilst wearing headphones to block out external sounds?

   Never 1 → 5 Frequently ___________

(FOCUS)

6. When playing video-games, do you make sure that the room is free of distractions before playing?

   Never 1 → 5 Frequently ___________

(FOCUS)

7. Do you find yourself becoming deeply involved in narrative mediums (movies, TV, books, theatre, video-games)?

   Never 1 → 5 Frequently ___________

(FOCUS)

8. Do you find that you lose track of time whilst playing video-games so much so that you become unaware of things happening around you?

   Never 1 → 5 Frequently __________

(FOCUS)

9. Do you ever become so engrossed in a video game that it feels like you are inside the game?

   Never 1 → 5 Frequently __________

(INVOL)

10. Do you ever become so involved in narrative mediums that other people have difficulty getting your attention?

    Never 1 → 5 Frequently __________

(GAMES)

11. If so, is there a certain style or genre of game that you notice this happening with most?

    ….. …………………………………………..

(INVOL)

12. Do you find yourself becoming emotionally involved with characters in a game and react to game events as they would?

    Never 1 → 5 Frequently __________

(GAMES)

13. If so, is there a certain style or genre of game that you notice this happening with most?

    …………………………………………..

(GAMES)

14. Could you please name your top three favourite games?

    …………………………………………..
    …………………………………………..
    …………………………………………..

Applicant name: Date:

Thank you for taking time to complete this form. *The test will form part of my MA thesis investigating immersion in video-games. The test will involve playing a third-person RPG, whilst being recorded, in a room that has been specifically set up for playing games. Afterwards, you will be asked to join the researcher and answer a few questions about your experience, whilst re-watching the captured footage. The experiment will be unpaid, but drinks and snacks will be provided before and during the experiment. All information collected will be strictly confidential and will not be shared outside of this experiment, though it may be stored for later use in the work.*

Appendix 4 – The Post-Gameplay Interview

Date:
Place:
Interviewer: Matthew Varley
Interviewee:

Icebreaker statement
1. **So, did you enjoy the game?**

---

2. **Was there anything in particular that you liked about it?**

---

3. **Was there anything external the game, such as physical controls, room, temperature, seating, that ruined the gameplay for you?**

---

4. Did you feel that the game was engaging and that you became involved in the narrative?

---

5. Did you find yourself becoming emotionally responsive to the character?

---

6. Were there any elements of the game itself that you think could be improved or that you found confusing?

---

7. Did you notice any changes to the sounds in the game as you were playing?

(*If audio changes were noticed, continue, if not, move to Question 10)*

8. Did the changes in audio have a negative impact on your gameplay experience?

---

9. If they hadn't have happened, do you think you would have been more immersed in the game-world?

(Move to video point 1 of audio changes in observation notes)

10. Did you notice that at this point in the game, the characters footsteps became out of sync?

If yes: did this have a negative impact on your gameplay experience?
If no: why do you think you didn't notice?

(Move to video point 2 of audio changes in observation notes)

11. Did you notice that at this point in the game, the characters footsteps didn't sound like footsteps?

If yes: did this have a negative impact on your gameplay experience?
If no: why do you think you didn't notice?

(Move to video point 3 of audio changes in observation notes)

12. Did you notice that at this point in the game, the characters footsteps were removed completely?

If yes: did this have a negative impact on your gameplay experience?
If no: why do you think you didn't notice?

(Move to video point 4 of audio changes in observation notes)

13. Did you notice that at this point in the game, although the character wasn't in a forest, the bird sounds continued?

If yes: did this have a negative impact on your gameplay experience?
If no: why do you think you didn't notice?

14. Can you think of any examples of sounds in games that have ruined your gaming experience in the past?

Thanks for your time, I hope you've had a good experience.

References

Anderson, J. D. (1996). *The reality of illusion: An ecological approach to cognitive film theory.* Carbondale: Southern Illinois University Press.

Andronico, M. (2016). Xbox Design Lab Review: Is This Custom Controller Worth It?. Retrieved from https://www.tomsguide.com/us/xbox-design-lab,review-3812.html

Bernstein, D. (1997). *Creating an interactive audio environment*. Retrieved from https://www.gamasutra.com/view/feature/131646/creating_an_interactive_audio_.php

Black, D. (2017). Why Can I See My Avatar? Embodied Visual Engagement in the Third-Person Video-Game. *Games and Culture, 12*(2). 179-199.

Boyd, A. (2003). When Worlds Collide: Sound and Music in Film and Games. Retrieved from https://www.gamasutra.com/view/feature/131310/when_worlds_collide_sound_and_.php

Bridgett, R. (2008). Dynamic range: subtlety and silence in video game sound. In Collins, K. (Eds.), *From Pac-Man to Pop Music: Interactive Audio in Games and New*

Media. (pp. 127-134). Hampshire, England: Ashgate Publishing Limited.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45*(2009). 624-634. doi:10.1016/j.jesp.2009.02.2016.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. Paper presented at the Conference for Human-Computer Interaction (CHI), Vienna, Austria. 1297-1300. doi:10.1145/985921.986048

Brown, M., Kehoe, A., Kirakowski, J., Pitt, I. (2011). Beyond the Gamepad: HCI and Game Controller Design and Evaluation. In Bernhaupt, R. (Eds.), *Evaluating user experience in games: Concepts and methods* (pp. 209-232). London: Springer. doi:10.1007/978-1-84882-963-3.

Calleja, G. (2011). *In-game: From immersion to incorporation.* US: MIT Press.

Cheng, K., & Cairns, P. (2005). Behaviour, Realism and Immersion in Games. Paper presented at the Conference for Human-Computer Interaction (CHI), Vienna, Austria. 1272-1275. doi:10.1145/1056808.1056894.

Cheng, W. (2014). *Sound Play: Video games and the Musical Imagination.* New York, USA: Oxford University Press.

Chion, M. (1994). *Audio-vision: Sound on Screen*. New York, USA: Columbia University Press.

Collins, K. (2007). *An Introduction to the Participatory and Non-Linear Aspects of Video Games Audio.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.85.7281&rep=rep1&type=pdf

Collins, K. (2008). Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design. Cambridge, England: MIT Press.

Cowley, B., Charles, D., Black, M., Hickey, R. (2008). Towards an Understanding of Flow in Video Games. *Computers in Entertainment*, 6(2). 1-27. doi:10.1145/1371216.1371223.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches (*Fourth, International student ed.). Los Angeles: SAGE.

Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience.* New York, USA: HarperPerennial.

Curtiss, S. (1992). The Sound of the Early Warner Bros. Cartoons'. In Altman, R. (Eds.). *Sound theory sound practice*. (191-203). London: Routledge.

Denisova, A., Cairns, P. (2015). First Person vs. Third Person Perspective in Digital Games: Do Player Preferences Affect Immersion?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 145-148). Seoul, Republic of Korea. doi: 10.1145/2702123.2702256.

Deutsch, S. (2011). From Battlezone to Battlefield Immersion, Sound and Computer Games: An Interview with Ben Minto. *The New Soundtrack,* 1.1(2011). 43-56. Doi:10.3366/sound.2011.0005.

DiGiose, N. 2016. *Surround sound vs. 3D sound*. Retrieved from: https://www.electronicproducts.com/News/Surround_sound_vs_3D_sound.aspx

Douglas, Y., Hargadon, A. (2000). The Pleasure Principle: Immersion, Engagement, Flow. Paper presented at the ACM on hypertext and hypermedia. 153-160. doi:10.1145/336296.336354.

Ekman, I. (2005). Meaningful Noise: Understanding Sound Effects in Computer Games. In *Proceedings of the 2005 conference on Digital Arts & Culture*. Kopenhagen, Denmark.

Ermi, L., Mayra, F. (2005, June). Fundamental Components of the Gameplay Experience: Analysing Immersion. In *Proceedings of Changing views: Worlds in Play*. DiGRA. Vancouver, Canada.

Fay, T. M., Selfon, S., Fay, T. J. (2004). DirectX 9 audio exposed: Interactive audio development. Plano, Tex: Wordware
Pub. Accessed via: http://library.books24x7.com.libaccess.hud.ac.uk/assetviewer.aspx?bookid=6829&chunkid=379783044&noteMenuToggle=0&leftMenuState=1

Frasca, G. (2001). Rethinking agency and immersion: video games as a means of consciousness-raising. *Digital Creativity, 12(*3). 167-174.

Gard, T. (2010). Action Adventure Level Design: Kung Fu Zombie Killer. Available at: https://www.gamasutra.com/view/feature/4413/action_adventure_level_design_.php

Geelen, T. V. (2008) Realizing groundbreaking adaptive music. In Collins, K. (Eds.), *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*. (pp. 94-102). Hampshire, England: Ashgate Publishing Limited.

Gorbman, C. (1987). *Unheard Melodies: Narrative Film Music*. Bloomington, Indiana University Press.

Grimshaw, M. (2007). *The Acoustic Ecology of the First-Person Shooter (Doctoral dissertation, The University of Waikato).* Retrieved from http://wikindx.com/mainsite/acousticEcologyOfTheFirst-PersonShooter.pdf

Grimshaw, M., Schott, G. (2008). A Conceptual Framework for the Analysis of First-Person Shooter Audio and its Potential Use for Game Engines. *International Journal of Computer Games Technology, 2008,* 1-7. doi:10.1155/2008/720280.

Grimshaw, M. (2012). Sound and Player Immersion in Digital Games. In Pinch, T., and Bijsterveld, K. (Eds.), *The Oxford Handbook of Sound Studies*. Oxford, England: Oxford University Press. doi:10.1093/oxfordhb/9780195388947.013.0084.

Guerraz. A., Lemordant, J. (2008). Indeterminate adaptive digital audio for games on mobiles. In Collins, K. (Eds.), *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*. (pp. 55-74). Hampshire, England: Ashgate Publishing Limited.

Hoffman, D. L., Novak, P. T. (2009). Flow Online: Lessons Learned and Future Prospects. *Journal of Interactive Marketing*, *23*(1). 23-34. doi:10.1016/j.intmar.2008.10.003.

Horowitz, S. (2012). *The Universal Sense: How Hearing Shapes the Mind.* New York, USA: Bloomsbury.

Hou, J., Nam, Y., Peng, W., Lee, K. M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior, 2*(2012). 617-623.

Ijsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: Concept, determinants, and measurement. Paper presented at the SPIE *3959*(1) doi:10.1117/12.387188

Ijsselsteijn, W. A., Freeman, J., & de Riddler, H. (2001). Presence, Where are we?. *CyberPsychology and Behavior*, 4. 179-182.

Jennett, C., Cox, L. A., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*(2008). 641-661.

Jin, S. A. (2012). "Toward Intergrative Models of Flow": Effects of Performance, Skill, Challenge, Playfulness, and Presence on Flow in Video Games. *Journal of Broadcasting & Electronic Media, 56*(2). 169-186. doi: 10.1080/08838151.

Jorgensen, K. (2008). Left in the dark: playing computer games with the sound turned off. In Collins, K. (Eds.), *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*. (pp. 163-176). Hampshire, England: Ashgate Publishing Limited.

Jorgensen, K. (2011). Time for New Terminology? Diegetic and Non-Diegetic Sounds in Computer Games Revisited. In Grimshaw, M. (Eds.), *Game Sound Technology and Player Interaction: Concepts and Developments.* (pp. 78-97). Hershey, PA: Information Science Reference.

Kamp, M. (2012). Musical Ecologies in Video Games. *Philosophy & Technology, 27*(2), 235-249. doi:10.1007/s13347-013-0113-z.

Kim, K., Schmierbach, M. G., Bellur, S., Chung, M., Fraustino, J. D., Dardis, F., Ahern, L. (2015). Is it a sense of autonomy, control, or attachment? Exploring the effects of in-game customization on game enjoyment. *Computers in Human Behavior, 48*(2015). 695-705. doi: 10.1016/j.chb.2015.02.011.

Kline, S., Dyer-Witheford, N., & de Peuter, G. (2003). *Digital play: The interaction of technology, culture, and marketing*. Montréal: McGill-Queen's University Press.

Laird, J. (2015). Projectors: A Guide to Truly Cinematic Gaming. Available at https://www.rockpapershotgun.com/2015/10/08/projectors-gaming-guide/

Laramée, F. D. (2002). *Game design perspectives*. Charles River Media: Massachusetts.

Lazzaro, N. (2004). *Why We Play Games: Four Keys to More Emotion Without Story.* Paper presented at the Game Developers Conference (GDC). Retrieved from http://xeodesign.com/xeodesign_whyweplaygames.pdf

Lebowitz, J., Klug, C. (2011). *Interactive Storytelling for Video Games*. Focal Press: Oxford.

Lemarchand, R. (2012) Attention, not immersion: Making Your Games Better with Psychology and Playtesting, the Uncharted Way [Powerpoint slides]. Retrieved from https://www.gdcvault.com/play/1015464/Attention-Not-Immersion-Making-Your

Liljedahl, M. (2011). Sound for Fantasy and Freedom. In Grimshaw, M. (Eds.), *Game Sound Technology and Player Interaction: Concepts and Developments.* (pp. 22-43). Hershey, PA: Information Science Reference.

Marks, A. (2009). *The complete guide to game audio: For composers, musicians, sound designers, game developers* (2nd ed.). Oxford: Focal. doi: 10.4324/9780080928074.

Martinez, M. A. (2014). Storyworld possible serlves and the pheneomenon of narrative immersion: Testing a new theortical construct. *Narrative, 22*(1), 110-131.

Madigan, J. (2010). *The Psychology of Immersion in Video* Games. Available at http://www.psychologyofgames.com/2010/07/the-psychology-of-immersion-in-video-games/#foot_text_514_2

McGloin, R., Farrar, K., Krcmar, M. (2013). Video Games, Immersion, and Cognitive Aggression: Does the Controller Matter?. *Media Psychology*, 16(1). 65-87.

McMahan, A. (2003). *Immersion, Engagement, and Presence; A New Method for Analysing 3-D Video Games*. Routledge, New York.

Morris, S. J. (2002). First Person Shooters: A Game Apparatus. In G. King and T. Krzywinska (Ed.), *ScreenPlay: Cinema/videogames/interfaces*. 81-97. London: Wallflower.

Moser, M. A., Macleod, D. (1996). Immersed in Technology: Art and Virtual Environments. MIT Press: Cambridge, MA.

Nacke, L. E., Lindley, C. A. (2008). Flow and Immersion in First-Person Shooter: Measuring the player's gameplay experience. Paper presented at the Conference on future play. 81-88. doi:10.1145/1496984.1496998.

Nacke, L. E., Grimshaw, N. M., Lindley, A. C. (2010). More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers*, 22(2010). 336-343. doi:10.1016/j.intcom.2010.04.005.

O'Keeffe, L. (2011). Sound is Not a Simulation: Methodologies for Examining the

Experience of Soundscapes. In Grimshaw, M. (Eds.), *Game Sound Technology and Player Interaction: Concepts and Developments.* (pp. 44-59). Hershey, PA: Information Science Reference.

Ortqvist, D., Liljedahl, M. (2010). Immersion and Gameplay Experience; A Contingency Framework*. International Journal of Computer Games Technology*, 2010. p.1-11

Patterson, N., Conway, F. (2014). Engagement, Immersion and Presence: The Role of Audio Interactivity in Location-aware Sound Design. In Collins, K., Kapralos, B., & Tessler, H. (Eds.). *The Oxford Handbook of Interactive Audio*. doi:1093/oxfordhb/9780199797226.013.008.

Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J. A., Verratti, T., Chiselko, G. (2000). Using a Large Projection Display Screen as an Alternative to Head-Mounted Displays for Virtual Environments. *CHI Letters,* 2(1). 478-485.

Porter, J. (2017). Gaming on a projector: the pros and cons of big-screen gaming. Available at: http://www.techradar.com/news/gaming-on-a-projector-the-pros-and-cons-of-big-screen-gaming

Qin, H., Patrick Rau, P., & Salvendy, G. (2009). Measuring Player Immersion in the Computer Game Narrative*. International Journal of Human–Computer Interaction*, *25*(2), 107-133. doi: 10.1080/10447310802546732.

Radford, A. (2000). Games and Learning about Form in Architechture. *Automation in Construction*, 9(2000). 379-385.

Roden, T. E., Parberry, I., Ducrest, D. (2005). Toward mobile entertainment: A paradigm for narrative-based audio only games. *Science of Computer Programming, 67*(2007) 76-90. doi: 10.1016/j.scico.2006.07.004.

Roux-Girard, G. (2014). Sound and the Videoludic Experience. In Collins, K., Kapralos, B., & Tessler, H. (Eds.). *The Oxford Handbook of Interactive Audio*. doi:1093/oxfordhb/9780199797226.013.008.

Ruddle, R., Payne, S., Jones, D. (1998). Navigating large-scale virtual environments: What differences occur between helmet-mounted and desktop displays. *Presence teleoperators and virtual environments*, 8(2). 157-168.

Ryan, M. (2001). *Narrative as virtual reality: Immersion and interactivity in literature and electronic media*. Baltimore, Maryland: The John Hopkins University Press.

Sanger, G. A. (2004). *The Fat Man on Game Audio: Tasty Morsels of Sonic Goodness*. USA: New Riders Publishing.

Schafer, M. R. (1994) *The Soundscape: Our Sonic Environment and the Tuning of the*

World. Rochester, Vermont: Destiny Book.

Schmierbach, M., Limperos, A. M., Woolley, J. K. (2012). Feeling the Need for (Personalized) Speed: How Natural Controls and Customization Contribute to Enjoyment of a Racing Game Through Enhanced Immersion. *CyberPsychology, Behavior, and Social Networking, 15*(7). 364-369. doi: 10.1089/cyber.2012.0025.

Sherry, J. L. (2004). Flow and Media Enjoyment. *Communication Theory, 14*(4). 328-347.

Smahel, D., Blinka, L., Ledabyl, O. (2008). Playing MMORPGs: Connections between Addiction and Identifying with a Character. *CyberPsychology & Behavior, 11*(6). 715-718. doi: 10.1089/cpb.2007.0210.

Somberg, G. (2016). *Game Audio Programming: Principles and Practices.* Milton: CRC Press Inc.

Sonnenschein, D. (2001). *Sound Design: The Expressive Power of Music, Voice, and Sound Effects in Cinema*. Studio City, CA: Michael Wiese Productions.

Stephanie. (2016). *Jaccard Index / Similarity Coefficient*. Available at http://www.statisticshowto.com/jaccard-index/

Stevens, R., & Raybould, D. (2015). The Reality Paradox: Authenticity, fidelity, and the real in Battlefield 4. *The Soundtrack, 8*(1-2). 57.75.

Stevens, R., & Raybould, D. (2016). *Game audio implementation: A practical guide to using the unreal* engine (Second ed.). Boca Raton: CRC Press. doi:10.4324/9781315772783.

Stockburger, A. (2003, November). The Game Environment from an Auditive Perspective. In *Proceedings of Level Up*. DiGRA. Utrecht Universiteit, Netherlands.

Suddaby, P. (2012). *The Many Ways to Show the Player How It's Done With In-Game Tutorials*. Retrieved from https://gamedevelopment.tutsplus.com/tutorials/the-many-ways-to-show-the-player-how-its-done-with-in-game-tutorials--gamedev-400

Sullivan, G. M. (2013). *Analyzing and Interpreting Data from Likert-Type Scales.* Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3886444/

Takatalo, J., Hakkinen, J., Kaistinen, J., & Nyman, G. (2010). Presence, Involvement, and Flow in Digital Games. In Bernhaupt, R. (Eds.), *Evaluating user experience in games: Concepts and methods* (pp. 23-46). London: Springer. doi:10.1007/978-1-84882-963-3.

Taylor, L. N. (2002). *Video Games: Perspective, Point-of-view, and immersion (master's* *thesis).* University of Florida, USA.

Taylor, G. (2010). *All in the Mix: The Importance of Real-Time Mixing in Video Games*. Paper presented at the Develop Conference. [Transcript]. Retrieved from http://gameaudionoise.blogspot.co.uk/p/all-in-mix-importance-of-real-time.html

Teng, C. (2010). Customization, immersion satisfaction, and online gamer loyalty. *Computers in Human Behavior, 26* (2010). 1547-1554.

Usher, R. (2012). *How Does In-Game Audio Affect Players?* Available at http://www.gamasutra.com/view/feature/168731/how_does_ingame_audio_affect_.php

Wengraf, T. (2001). *Qualitative Research Interviewing*. London: Sage Publications.

Westerberg, A., Schoenau-Fog, H. (2015). Categorizing Video Game Audio – An exploration of auditory-psychological effects. Paper presented at MINDTREK. 47-54. doi:10.1145/2818187.2818295.

Whalen, Z. (2004). *Play Along – An Approach to Videogame Music*. Retrieved from: http://www.gamestudies.org/0401/whalen/

Wissmath, B., Weibel, D., Groner, R. (2009). Dubbing or Subtitling? Effects on Spatial Presence, Transportation, Flow, and Enjoyment. *Journal of Media Psychology,* 21(3). 114-125.

Witmer, B. G., Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3). 225–240.

Wu, H., Guo, Y., & Seaman, C. B. (2009). Analyzing video data: A study of programming behavior under two software engineering paradigms. Paper presented at the 456-459. doi:10.1109/ESEM.2009.5315995.

Videography

FKN Ensemble. (2016). *Top 10 Most Immersive Game Worlds* [Video File]. Retrieved from https://www.youtube.com/watch?v=LB6HZAlj_pU&index=3&list=PLyMZ3Sa3M9ix-NSaJzFUI5o7ESWXp1Riw

Frantzolas, T. [TEDx Talks]. (2016). *The Beautiful Lies of Sound Design | Tasos Fratzolas | TEDxAthens* [Video File]. Retrieved from https://www.youtube.com/watch?v=jDy5j0c6TrU&t=22s&list=PLyMZ3Sa3M9izPCmHzx5Tn43Ey-NV4xJCK&index=46

Kramer, B. [GrumpOut]. (2016). *What Makes a Game Immersive? – HOW ABOUT THIS GAME? – GrumpOut* [Video File]. Retrieved from https://www.youtube.com/watch?v=Xncs1mrMV6o&list=PLyMZ3Sa3M9ix-NSaJzFUI5o7ESWXp1Riw&index=4

Perryman, M. [Michael Perryman]. (2016). *An Insight to Gaming – Sound Design* [Video File]. Retrieved from https://www.youtube.com/watch?v=hO6QcN9bH-c&t=54s

TenFortySeven. (2016). *What is Immersion? – Immersion Gaming #1* [Video File]. Retrieved from https://www.youtube.com/watch?v=sfm1RjvrK6k&list=PLyMZ3Sa3M9ix-NSaJzFUI5o7ESWXp1Riw&index=2

Soundworks Collection. (2016a). *The Sound of Uncharted 4: A Thief's End* [Video File]. Retrieved from http://soundworkscollection.com/videos/the-sound-of-uncharted-4-a-thief-s-end

Soundworks Collection. (2016b). *The Music of Far Cry Primal with Composer Jason Graves* [Video File]. http://soundworkscollection.com/videos/the-music-of-far-cry-primal-with-composer-jason-graves

Soundworks Collection. (2016c). *The Sound and Music of The Last of Us* [Video File]. Retrieved from http://soundworkscollection.com/videos/the-sound-and-music-of-the-last-of-us

Soundworks Collection. (2016e). *The Sound and Music of HALO: REACH* [Video File]. Retrieved from http://soundworkscollection.com/videos/halo-reach

Steinberg. (2016). *Game audio design with Nuendo for Assassin's Creed IV Black Flag* [Video File]. Retrieved from https://www.youtube.com/watch?v=QULt-fp-r-A&feature=youtu.be

The Creators Project. (2013). *The Terrifying Sound Design of Dead Space 3: Kill Screen Episode 5* [Video File]. Retrieved from https://www.youtube.com/watch?v=pv8mkxl8fbs

Woelfer, A. [Video Game Close Up]. (2015). *What is immersion? | Game Studies* [Video File]. Retrieved from https://www.youtube.com/watch?v=XpZNJ-S5U30&index=1&list=PLyMZ3Sa3M9ix-NSaJzFUI5o7ESWXp1Riw

Gameography

Blizzard Entertainment. (2005). *World of Warcraft*. EU: Blizzard Entertainment.

Core Design. (1996). *Tomb Raider*. EU: Eidos Interactive.

Genki. (2000). *Jade Cocoon: Story of the Tamamayu*. EU: Crave Entertainment.

Genki. (2002). Jade Cocoon 2. EU: Ubi Soft Entertainment.

Monolith Productions. (2014). *Middle-earth: Shadow of Mordor*. EU: Warner Bros. Interactive Entertainment.

Rockstar North. (2013). *Grand Theft Auto V*. WW: Rockstar Games.

Square. (1987). *Final Fantasy*. JP: Square.